

Higher Mathematics

Ya.S. BUGROV S.M. NIKOLSKY

**fundamentals
of
linear algebra
and
analytical
geometry**

MIR PUBLISHERS MOSCOW

Higher Mathematics

Higher Mathematics

Ya. S. BUGROV

S. M. NIKOLSKY

**fundamentals
of
linear algebra
and
analytical
geometry**

Translated from the Russian
by
Leonid Levant

MIR PUBLISHERS
MOSCOW

First published 1982
Revised from the 1980 Russian edition

На английском языке

- © Издательство «Наука», Главная редакция физико-математической литературы, 1980
- © English translation, Mir Publishers, 1982

Contents

Preface	6
Sec. 1. Second-order Determinants	7
Sec. 2. Determinants of the Third and n th Order	8
Sec. 3. Matrices	19
Sec. 4. Systems of Linear Equations. The Kronecker-Capelli Theory	21
Sec. 5. Three-dimensional Space. Vectors. Cartesian System of Coordinates	39
Sec. 6. n -Dimensional Euclidean Space. Scalar Product of Two Vectors	48
Sec. 7. Segment of a Line. Dividing a Segment in a Given Ratio	53
Sec. 8. The Straight Line	56
Sec. 9. The Equation of a Plane	66
Sec. 10. A Straight Line in Space	74
Sec. 11. Orientation of Rectangular Coordinate Systems	78
Sec. 12. The Vector Product of Two Vectors	81
Sec. 13. Triple Scalar Product	87
Sec. 14. Linearly Independent System of Vectors	89
Sec. 15. Linear Operators	96
Sec. 16. Bases in R_n	102
Sec. 17. Orthogonal Bases in R_n	107
Sec. 18. Invariant Properties of a Scalar and a Vector Product	113
Sec. 19. Transformation of Rectangular Coordinates in a Plane	116
Sec. 20. Linear Subspaces in R_n	119
Sec. 21. Fredholm-type Theorems	125
Sec. 22. Self-adjoint Operator. Quadratic Form	132
Sec. 23. Quadratic Form in Two-dimensional Space	142
Sec. 24. Second-order Curves	146
Sec. 25. Second-order Surfaces in Three-dimensional Space	162
Sec. 26. The General Theory of a Second-order Surface in Three-dimensional Space	180
Subject Index	187

Preface

The present book is the first part of our three-part textbook "Higher Mathematics". We deal here with the fundamentals of the theory of determinants, the elements of the theory of matrices, the theory of systems of linear equations, and vector algebra. The book is also intended to introduce its readers to the basic aspects of linear algebra: linear operators, orthogonal transformations, self-adjoint operators, the quadratic form and reducing it to the canonical form. Elements of analytical geometry (the straight line, the plane, the straight line in space, and second-order curves and surfaces) are also included.

As a rule, our arguments are accompanied by exhaustive proofs. The material is introduced so that the proofs for the general n -dimensional case may be omitted without the loss of either the formulation of the statement or a detailed explanation of how the subject matter stands in the two- and three-dimensional cases.

The canonical forms of second-order curves and surfaces are treated briefly in this book since it is assumed that they will be studied additionally by solving particular problems using the methods of mathematical analysis. The quadratic form is dealt with by the methods of mathematical, or, if more convenient, by functional analysis.

Although we have mentioned that this is the first book in our series, its material is closely interwoven with that of the second book dedicated to differential and integral calculus. The third book will consider differential equations, multiple integrals, series, and the theory of functions of a complex variable. Being complementary books of a single series entitled "Higher Mathematics", at the same time they may be regarded as independent educational aids for students of higher engineering institutions.

The authors

Sec. 1. Second-order Determinants

Let there be given numbers a_1, a_2, b_1, b_2 . They define the number $a_1b_2 - a_2b_1$ which is called the *determinant of the second order* and is written in the following way:

$$\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} = a_1b_2 - a_2b_1. \quad (1)$$

The numbers a_1, a_2, b_1, b_2 are termed *elements of the determinant*. Considering determinant (1), we distinguish the *first row* a_1, a_2 and the *second row* b_1, b_2 , the *first column* a_1, b_1 and the *second column* a_2, b_2 .

One can easily verify the following *properties of determinants*.

The value of a determinant:

(a) *does not change if each of the rows is substituted by a column of the same position number:*

$$\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix};$$

(b) *changes sign if its rows or columns are interchanged:*

$$\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} = - \begin{vmatrix} b_1 & b_2 \\ a_1 & a_2 \end{vmatrix},$$

$$\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} = - \begin{vmatrix} a_2 & a_1 \\ b_2 & b_1 \end{vmatrix};$$

(c) *increases k times if the elements of any of its columns or rows are increased k times. For instance:*

$$\begin{vmatrix} ka_1 & b_1 \\ ka_2 & b_2 \end{vmatrix} = k \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix},$$

i.e. a common factor of all the elements of one row (or column) may be taken outside the sign of the determinant;

(d) is equal to zero if the elements of a certain column or row are equal to zero. For instance:

$$\begin{vmatrix} 0 & 0 \\ b_1 & b_2 \end{vmatrix} = 0b_2 - 0b_1 = 0;$$

(e) is equal to zero if the determinant has two identical rows or columns:

$$\begin{vmatrix} a_1 & a_1 \\ a_2 & a_2 \end{vmatrix} = a_1a_2 - a_1a_2 = 0.$$

Introduced below are determinants of the third and, in general, of the n th order. For them the above listed properties hold true.

Sec. 2. Determinants of the Third and n th Order

The number

$$\begin{aligned} \Delta = & a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ & - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}, \end{aligned} \quad (1)$$

written in the form

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}, \quad (2)$$

is called the *determinant of the third order*.

In determinant (2) we distinguish the first, second, and third rows, as well as the first, second, and third columns. The number a_{kl} is called the *element* of the determinant in row k and column l , where k is the number of the row and l is the number of the column. The element a_{kl} is said to be found at the intersection of the k th row and l th column.

The number of rows (or columns) is the order of the determinant. The elements a_{11} , a_{22} , a_{33} form the *principal* (or *leading*) *diagonal* of the determinant, and the elements a_{13} , a_{22} , a_{31} form the *secondary* diagonal. In other words: the diagonal from the upper left corner to the lower right

corner (on which the elements a_{11} , a_{22} , a_{33} are located) is called the principal diagonal of determinant (2).

The structure of expression (1) is rather simple. This is a number expressed in terms of elements a_{hl} and computed according to Sarrus' rule. Shown below (see Fig. 1) is a Sarrus table compiled from the elements of determinant (2) by adding the first and second columns of the

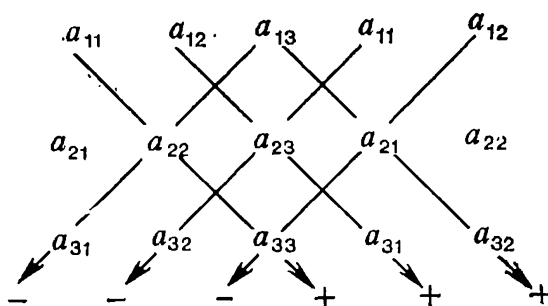


Fig. 1.

determinant. As is obvious from the figure, we have to take all possible products of the elements "crossed out" by the straight lines. In doing so, the three products indicated by the lines parallel to the principal diagonal should be taken with the plus sign, and the remaining three products corresponding to the lines parallel to the secondary diagonal are taken with the minus sign. This is just Sarrus' rule for determining the signs of terms in the expansion of a determinant.

Every product preceded by a certain sign is called the *term of determinant* (2). Among the elements entering into these products there are "representatives" from each row and from each column. In each term these elements can be arranged in the increasing order of the first index, i.e. the index of the row they belong to. This is just done in sum (1). As far as the numbers of the columns to which these elements belong are concerned, their arrangements are given below:

$$\left. \begin{array}{l} 1 \ 2 \ 3, \\ 2 \ 3 \ 1, \\ 3 \ 1 \ 2, \end{array} \right\} \quad (3)$$

$$\left. \begin{array}{ccc} 3 & 2 & 1, \\ 1 & 3 & 2, \\ 2 & 1 & 3. \end{array} \right\} \quad (4)$$

Obviously, these are all possible permutations of the digits 1, 2, 3. The permutation

$$1, 2, 3 \quad (5)$$

will be called the *main* permutation.

We speak of a *transposition of two certain elements* in a permutation if their places are interchanged. After this transposition the permutation turns into another permutation which, in its turn, after a new transposition yields a third permutation (possibly, the first one). For instance, the permutation

$$3, \quad 2, \quad 1 \quad (6)$$

is obtained by transposing the first and third elements of permutation (5), and the permutation

$$2, \quad 3, \quad 1 \quad (7)$$

by transposing the first and second elements of permutation (6).

It is important to note that if a certain permutation is obtained from the main one via N transpositions and if the same permutation is obtained in some other way from the main transposition by means of N_1 transpositions, then both numbers N and N_1 are simultaneously either even, or odd. The permutation of the digits 1, 2, 3 is said to be *even (odd)* if it is obtained from the main permutation via an even (odd) number of transpositions.

Let there be given a permutation $j = (j_1, j_2, j_3)$, where j_1, j_2, j_3 are the numbers 1, 2, 3 arranged in a certain order. Let us denote by $t(j)$ the *number of transpositions* by means of which it is possible to get this permutation from the main one. Then the permutation j is even (odd) if $t(j)$ is an even (odd) number.

Permutations (3) are even, and (4) are odd.

According to what has been said, we may give another (equivalent) definition of a third-order determinant,

The *determinant of the third order* (2) is defined as a number Δ equal to the sum

$$\Delta = \sum_j (-1)^{t(j)} a_{1j_1} a_{2j_2} a_{3j_3} \quad (8)$$

of products of the form $(-1)^{t(j)} a_{1j_1} a_{2j_2} a_{3j_3}$, where $j = (j_1, j_2, j_3)$ are all possible different permutations derived from the main permutation 1, 2, 3.

This definition is generalized for determinants of the n th order ($n = 1, 2, 3, \dots$).

The *determinant of the n th order* is a number written in the form

$$\Delta = |a_{ik}| = \begin{vmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} \quad (9)$$

and evaluated by the given numbers a_{ik} (elements of the determinant) according to the following law: Δ is the sum

$$\Delta = \sum_j (-1)^{t(j)} a_{1j_1} a_{2j_2} \dots a_{nj_n},$$

extended to all possible permutations $j = (j_1, \dots, j_n)$. The number $t(j)$ equals the *number of transpositions* which has to be done in order to pass from the main permutation 1, 2, $\dots$, n to the permutation $j = (j_1, \dots, j_n)$. The product $(-1)^{t(j)} a_{1j_1} \dots a_{nj_n}$ is called the *term of a determinant*.

Determinants of the n th order possess the properties (a), (b), (c), (d), and (e) discussed in the preceding section.

Proof. (a) After certain rows of a determinant are replaced by the corresponding columns, the row numbers will be denoted by the second indices. For instance, for third-order determinant (2) we shall have

$$\begin{vmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{21}a_{23}$$

$$- a_{31}a_{22}a_{13} - a_{11}a_{32}a_{23} - a_{21}a_{12}a_{33} = \Delta.$$

In the general case, the general term of the new determinant will be written as follows

$$(-1)^{t(j)} a_{j_1 1} a_{j_2 2} \dots a_{j_n n}.$$

Now we are going to order the factors of the product $a_{j_1 1} a_{j_2 2} \dots a_{j_n n}$ with respect to the first index, i.e. we pass over from the permutation $j = (j_1, \dots, j_n)$ to the main permutation $1, 2, \dots, n$. In doing so, we have to accomplish $t(j)$ transpositions. Then the main permutation of the second indices will turn into a certain permutation $i = (i_1, \dots, i_n)$ and the number $t(i)$ will be of the same parity as the number $t(j)$. Thus,

$$(-1)^{t(j)} a_{j_1 1} \dots a_{j_n n} = (-1)^{t(i)} a_{1 i_1} \dots a_{n i_n}.$$

It is not difficult to see that to different permutations $j_1, \dots, j_n$ there correspond different permutations $i_1, \dots, i_n$. But then

$$\sum_j (-1)^{t(j)} a_{j_1 1} \dots a_{j_n n} = \sum_i (-1)^{t(i)} a_{1 i_1} \dots a_{n i_n} = \Delta.$$

(b) Let us interchange, for instance, the first and third rows of third-order determinant (2). The newly obtained determinant denoted by Δ' will be equal to

$$\begin{aligned} \Delta' &= \begin{vmatrix} a_{31} & a_{32} & a_{33} \\ a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \end{vmatrix} = \sum_j (-1)^{t(j)} a_{3 j_1} a_{2 j_2} a_{1 j_3} \\ &= \sum_j (-1)^{t(j)} a_{1 j_3} a_{2 j_2} a_{3 j_1} \\ &= - \sum_{j'=(j_3, j_2, j_1)} (-1)^{t(j')} a_{1 j_3} a_{2 j_2} a_{3 j_1} = -\Delta, \end{aligned}$$

since the permutation $j = (j_1, j_2, j_3)$ differs from the permutation $j' = (j_3, j_2, j_1)$ by one transposition.

(c) Multiplication of some row (column) of a determinant by number k is reduced to multiplication of all its terms by k , since every term contains one element of the indicated row (column). But then the sum of the terms will be multiplied by k ,

(d) A determinant, the elements of some column or row of which are equal to zero, is zero, since all of its terms are, obviously, equal to zero.

(e) A determinant with two identical rows or two identical columns is equal to zero. This follows from property (b) ($\Delta' = -\Delta$, $\Delta' = \Delta$, where $\Delta = 0$).

In determinant (9) of the n th order let us delete the i th row and k th column. The remaining expression yields a determinant M_{ik} of the $(n - 1)$ st order which is called the *minor of the element* a_{ik} . And the quantity

$$A_{ik} = (-1)^{i+k} M_{ik}$$

is called the *cofactor of the element* a_{ik} .

Property (f). *The sum of the products of elements a_{ik} of some row (column) of a determinant by the cofactors of these elements is equal to the value of the determinant:*

$$\Delta = \sum_{k=1}^n a_{ik} A_{ik} \quad (i = 1, \dots, n), \quad (10)$$

$$\Delta = \sum_{i=1}^n a_{ik} A_{ik} \quad (k = 1, \dots, n). \quad (10')$$

Let us prove this property for a third-order determinant with respect to the third row. We have

$$\begin{aligned} & a_{31}A_{31} + a_{32}A_{32} + a_{33}A_{33} \\ &= a_{31} \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} - a_{32} \begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix} + a_{33} \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \\ &= a_{31} (a_{12}a_{23} - a_{13}a_{22}) + a_{32} (a_{13}a_{21} - a_{11}a_{23}) \\ &\quad + a_{33} (a_{11}a_{22} - a_{12}a_{21}) = \Delta. \end{aligned}$$

Sum (10) is called the *expansion of the determinant in terms of the elements of the i th row*, and sum (10') the *expansion of the determinant in terms of the elements of the k th column*.

Example 1. If the determinant Δ (see (9)) $a_{21} = a_{31} = \dots = a_{n1} = 0$, then $\Delta = a_{11}A_{11}$, i.e. evaluation of this determinant is reduced to evaluation of one of its cofactors, i.e. of a determinant of the $(n - 1)$ st order.

Example 2. If all the elements of Δ standing below (above) the principal diagonal of Δ are equal to zero ($a_{kl} = 0$, if $k > l$ ($k < l$)), then $\Delta = a_{11}a_{22} \dots a_{nn}$. This follows from the preceding example.

Property (g). *The sum of products of the elements a_{ik} of some row (column) of determinant by the corresponding cofactors of the elements of another row (column) is equal to zero:*

$$\sum_{k=1}^n a_{ik}A_{jk} = \sum_{k=1}^n a_{ki}A_{kj} = 0 \quad (11)$$

($i \neq j$; $i, j = 1, \dots, n$).

Indeed, let us draw our attention to the first sum. This sum is independent of the elements of the j th row. Let us now replace the elements of the j th row by the corresponding elements of the i th row. This substitution leaves the sum under consideration unaltered. Meanwhile, it can be regarded as the expansion of the new determinant in terms of the elements of the j th row, and as such, it then equals the value of the new determinant. But the latter, by property (e), is zero, since it has two identical rows (i th and j th).

Property (h). *Let there be given two n th-order determinants Δ_1 and Δ_2 in which all the rows (columns) are identical, except for a certain one. The sum of such determinants is equal to a determinant Δ of the n th order in which the indicated row (column) consists of the sum of the corresponding elements of this row (column) of the determinants Δ_1 and Δ_2 . For instance,*

$$\begin{aligned} \Delta_1 + \Delta_2 &= \begin{vmatrix} a_{11} & \dots & a_{1, n-1} & a_{1n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & \dots & a_{n, n-1} & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & \dots & a_{1, n-1} & b_{1n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & \dots & a_{n, n-1} & b_{nn} \end{vmatrix} \\ &= \begin{vmatrix} a_{11} & \dots & a_{1, n-1} & (a_{1n} + b_{1n}) \\ \dots & \dots & \dots & \dots \\ a_{n1} & \dots & a_{n, n-1} & (a_{nn} + b_{nn}) \end{vmatrix} = \Delta. \end{aligned}$$

Indeed, expanding the given determinants in terms of the elements of the n th column, we get

$$\begin{aligned}\Delta_1 + \Delta_2 &= \sum_{k=1}^n a_{kn} A_{kn} + \sum_{k=1}^n b_{kn} A_{kn} \\ &= \sum_{k=1}^n (a_{kn} + b_{kn}) A_{kn} = \Delta.\end{aligned}$$

Property (i). *If to the elements of some row (column) we add the corresponding elements of another row (column) multiplied by number k , then the value of a determinant remains unchanged.* For instance,

$$\begin{aligned}& \begin{vmatrix} a_{11} & \dots & a_{1,n-1} & a_{1n} + ka_{11} \\ \dots & \dots & \dots & \dots \\ a_{n1} & \dots & a_{n,n-1} & a_{nn} + ka_{n1} \end{vmatrix} \\ &= \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} + k \begin{vmatrix} a_{11} & \dots & a_{1,n-1} & a_{11} \\ \dots & \dots & \dots & \dots \\ a_{n1} & \dots & a_{n,n-1} & a_{n1} \end{vmatrix} \\ &= |a_{kl}| + k \cdot 0 = |a_{kl}| \end{aligned}$$

by virtue of properties (h), (c), and (e).

Being properly applied, this property reduced evaluation of the given determinant to evaluation of a determinant of a lower order.

Example 3.

$$\begin{aligned}& \begin{vmatrix} 1 & 4 & 5 \\ 2 & 3 & 1 \\ 8 & 1 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 4 & 5 \\ 0 & -5 & -9 \\ 8 & 1 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 4 & 5 \\ 0 & -5 & -9 \\ 0 & -31 & -39 \end{vmatrix} \\ &= \begin{vmatrix} -5 & -9 \\ -31 & -39 \end{vmatrix} = \begin{vmatrix} 5 & 9 \\ 31 & 39 \end{vmatrix} = 5 \cdot 39 - 9 \cdot 31 = -84.\end{aligned}$$

Example 4.

$$\begin{vmatrix} 2 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 4 & 6 \end{vmatrix} = \begin{vmatrix} 0 & 1 & 0 \\ -5 & 3 & -1 \\ -7 & 4 & 2 \end{vmatrix} = - \begin{vmatrix} -5 & -1 \\ -7 & 2 \end{vmatrix} = \begin{vmatrix} 5 & 1 \\ -7 & 2 \end{vmatrix} = 17.$$

Example 5. The determinant

$$\Delta_n = \begin{vmatrix} 1 & a_1 & a_1^2 & \dots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \dots & a_2^{n-1} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & a_n & a_n^2 & \dots & a_n^{n-1} \end{vmatrix},$$

generated by the numbers $a_1, a_2, \dots, a_n$ is called a *power determinant* or *Vandermonde determinant**.

This determinant is equal to zero if any two numbers a_j and a_k are equal to each other. If all a_j 's are different, then

$$\begin{aligned} \Delta_n = & (a_n - a_1)(a_{n-1} - a_1) \dots (a_2 - a_1) \cdot \\ & \cdot (a_n - a_2)(a_{n-1} - a_2) \dots (a_3 - a_2) \cdot \\ & \dots \cdot (a_n - a_{n-2})(a_{n-1} - a_{n-2}) \cdot \\ & \cdot (a_n - a_{n-1}). \end{aligned} \quad (12)$$

Indeed, for $n=2$

$$\Delta_2 = \begin{vmatrix} 1 & a_1 \\ 1 & a_2 \end{vmatrix} = a_2 - a_1,$$

i.e. formula (12) is true. Let us assume that formula (12) also holds true for $n = k - 1$, and prove that it is true for $n = k$. In doing so, we shall take advantage of properties (i) and (c) of determinants. Multiplying the $(k - 1)$ th column in the determinant Δ_k by a_1 and subtracting it from the k th column, multiplying the $(k - 2)$ th column also by a_1 and subtracting it from the $(k - 1)$ th column, and so on, we then obtain:

$$\Delta_k = \begin{vmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & a_2 - a_1 & a_2(a_2 - a_1) & \dots & a_2^{k-2}(a_2 - a_1) \\ \dots & \dots & \dots & \dots & \dots \\ 1 & a_k - a_1 & a_k(a_k - a_1) & \dots & a_k^{k-2}(a_k - a_1) \end{vmatrix}$$

* Vandermonde, Alexandre Théophile (1735-1796). French algebraist. Gave first logical exposition of the theory of determinants.

$$= (a_k - a_1) (a_{k-1} - a_1) \dots (a_2 - a_1) \begin{vmatrix} 1 & a_2 & \dots & a_2^{k-2} \\ 1 & a_3 & \dots & a_3^{k-2} \\ \dots & \dots & \dots & \dots \\ 1 & a_k & \dots & a_k^{k-2} \end{vmatrix}.$$

The latter is also a Vandermonde determinant of order $(k - 1)$, generated by the numbers $a_2, \dots, a_k$, therefore, by the assumption, we have

[illegible]

Thus, by the method of mathematical induction, formula (12) holds true for any $n \geq 2$.

Property (j) *Let*

$$\Delta_1 = |b_{kl}|, \quad \Delta_2 = |a_{kl}|.$$

The product

$$\Delta_1 \Delta_2 = \left| \sum_{i=1}^n b_{kj} a_{jl} \right| = \Delta,$$

i.e. the product of two n th-order determinants with the elements b_{kl}, a_{kl} is also a determinant of the n th order with the elements

$$\gamma_{kl} = \sum_{j=1}^n b_{kj} a_{jl}.$$

Thus, the element γ_{kl} belonging to the k th row and l th column of the determinant Δ is said to be equal to the product of the k th row of the determinant Δ_1 by the l th column of the determinant Δ_2 . Actually, it is the sum of products of the elements of the k th row of the determinant Δ_1 by the corresponding elements of the l th column of the determinant Δ_2 .

Since in the determinants Δ_1 and Δ_2 rows and columns may be interchanged, obviously, the elements γ_{kl} of the

product Δ can also be constructed by taking the product of the k th row of Δ_1 by the l th row of Δ_2 , or the product of the k th column of Δ_1 by l th column of Δ_2 , or of the k th column by the l th row.

Proof. Let us make sure that this property is correct by considering a particular example of second-order determinants:

$$\Delta_1 = \begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{vmatrix}, \quad \Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \quad \Delta = \begin{vmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{vmatrix},$$

where

$$\begin{aligned} \gamma_{11} &= b_{11}a_{11} + b_{12}a_{21}, & \gamma_{12} &= b_{11}a_{12} + b_{12}a_{22}, \\ \gamma_{21} &= b_{21}a_{11} + b_{22}a_{21}, & \gamma_{22} &= b_{21}a_{12} + b_{22}a_{22}. \end{aligned}$$

By properties (h), (c), and (e), we have

$$\begin{aligned} \Delta &= \begin{vmatrix} b_{11}a_{11} & b_{11}a_{12} + b_{12}a_{22} \\ b_{21}a_{11} & b_{21}a_{12} + b_{22}a_{22} \end{vmatrix} + \begin{vmatrix} b_{12}a_{21} & b_{11}a_{12} + b_{12}a_{22} \\ b_{22}a_{21} & b_{21}a_{12} + b_{22}a_{22} \end{vmatrix} \\ &= a_{11}a_{12} \begin{vmatrix} b_{11} & b_{11} \\ b_{21} & b_{21} \end{vmatrix} + a_{11}a_{22} \begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{vmatrix} + a_{12}a_{21} \begin{vmatrix} b_{12} & b_{11} \\ b_{22} & b_{21} \end{vmatrix} \\ &\quad + a_{21}a_{22} \begin{vmatrix} b_{12} & b_{12} \\ b_{22} & b_{22} \end{vmatrix} = a_{11}a_{12} \cdot 0 + a_{11}a_{22}\Delta_1 + a_{12}a_{21} \begin{vmatrix} b_{12} & b_{11} \\ b_{22} & b_{21} \end{vmatrix} \\ &\quad + a_{21}a_{22} \cdot 0 = a_{11}a_{22}\Delta_1 - a_{12}a_{21}\Delta_1 \\ &= \Delta_1 (a_{11}a_{22} - a_{12}a_{21}) = \Delta_1\Delta_2. \end{aligned}$$

In the general case of n th-order determinants it is possible to write

$$\Delta = \begin{vmatrix} \sum_{s_1=1}^n a_{1s_1}b_{s_11} & \sum_{s_1=1}^n a_{1s_1}b_{s_12} & \dots & \sum_{s_1=1}^n a_{1s_1}b_{s_1n} \\ \sum_{s_2=1}^n a_{2s_2}b_{s_21} & \sum_{s_2=1}^n a_{2s_2}b_{s_22} & \dots & \sum_{s_2=1}^n a_{2s_2}b_{s_2n} \\ \dots & \dots & \dots & \dots \\ \sum_{s_n=1}^n a_{nsn}b_{s_n1} & \sum_{s_n=1}^n a_{nsn}b_{s_n2} & \dots & \sum_{s_n=1}^n a_{nsn}b_{s_nn} \end{vmatrix}$$

$$\begin{aligned}
 &= \sum_{s_1=1}^n \sum_{s_2=1}^n \dots \sum_{s_n=1}^n a_{1s_1} a_{2s_2} \dots a_{ns_n} \begin{vmatrix} b_{s_1 1} & b_{s_1 2} & \dots & b_{s_1 n} \\ b_{s_2 1} & b_{s_2 2} & \dots & b_{s_2 n} \\ \dots & \dots & \dots & \dots \\ b_{s_n 1} & b_{s_n 2} & \dots & b_{s_n n} \end{vmatrix} \\
 &= \sum_{s=(s_1, s_2, \dots, s_n)} a_{1s_1} a_{2s_2} \dots a_{ns_n} (-1)^{t(s)} \Delta_1 = \Delta_2 \Delta_1.
 \end{aligned}$$

In computing individual elements of Δ we have the right to choose any summation index s ($\gamma_{kl} = \sum_{s=1}^n a_{ks} b_{sl}$), but for the further reasoning it is convenient to take s_1 as such an index for the first row of Δ , s_2 for the second row, and so on. The second equality is due to properties (h) and (c), the multiple sum $\sum_{s_1=1}^n \sum_{s_2=1}^n \dots \sum_{s_n=1}^n$ being extended to all possible permutations $(s_1, s_2, \dots, s_n)$, where $1 \leq s_j \leq n$. But if in some system $(s_1, s_2, \dots, s_n)$ two components s_i and s_j are equal to each other ($s_i = s_j$, $i \neq j$), then the neighbouring determinant $|b_{s_k l}| = 0$. Therefore, actually, in the multiple sum we may leave only the terms corresponding to different permutations $(s_1, \dots, s_n)$ from natural numbers $(1, \dots, n)$. And, obviously, it will turn out that the corresponding determinant

$$|b_{s_k l}| = (-1)^{t(s)} \Delta_1.$$

Sec. 3. Matrices

A rectangular array of numbers α_{ij} of the form

$$A = \begin{vmatrix} \alpha_{11} & \dots & \alpha_{1n} \\ \alpha_{m1} & \dots & \alpha_{mn} \end{vmatrix} = \begin{pmatrix} \alpha_{11} & \dots & \alpha_{1n} \\ \alpha_{m1} & \dots & \alpha_{mn} \end{pmatrix} = \|\alpha_{ij}\| = (\alpha_{ij}), \quad (1)$$

consisting of m rows and n columns is called the *matrix*. The numbers α_{ij} are called its *elements*. If $m = n$, then we have a *square matrix* of the n th order or dimension. In other words, a square matrix is a matrix for which the number of rows is equal to the number of columns.

The order of a square matrix is simply the number of rows (columns) in the matrix.

If a second matrix is given: $B = \|\beta_{ij}\|$ with elements β_{ij} also consisting of m rows and n columns, then it is regarded as *equal* to the matrix A , if and only if the corresponding elements of both matrices are equal ($\alpha_{ij} = \beta_{ij}$). In this case we write $A = B$. The matrix $\|\alpha_{ij}\|$ is not a number—this is an array. But for a square matrix it is possible to consider the number $|\alpha_{ij}|$ which is a *determinant generated by this matrix*.

Let k be a natural number not exceeding m and n ($k \leq m, n$). Let us now cross out some k columns and k rows in array (1). Then the elements α_{ij} found at the intersection of the deleted columns and rows form a square matrix which generates a determinant of order k . The determinant thus obtained is called the *kth-order determinant generated by matrix A*.

The greatest natural number k for which there exists a nonzero determinant of the k th order generated by matrix A is called the *rank* of matrix A (see Note 2 of the next section).

If in matrix A its rows and columns are interchanged (leaving the same numbers), then we obtain the following matrix

$$A^* = \left\| \begin{array}{cccc} \alpha_{11} & \dots & \alpha_{m1} \\ \vdots & \ddots & \vdots \\ \alpha_{1n} & \dots & \alpha_{mn} \end{array} \right\|,$$

which is called a *conjugate matrix* or a *transpose of matrix A*.

Matrices of one and the same dimension, i.e. containing one and the same number of rows and columns, can be added. The *sum* of two such matrices $A = \|\alpha_{ij}\|$ and $B = \|\beta_{ij}\|$ is a matrix $C = \|\gamma_{ij}\|$, whose elements are equal to the sum of the corresponding elements of the matrices A and B : $\gamma_{ij} = \alpha_{ij} + \beta_{ij}$. Symbolically, we shall write this fact in the following way:

$$A + B = C.$$

It is easy to see that

$$A + B = B + A,$$

$$(A + B) + C = A + (B + C).$$

Example. Let

Find the matrix $\lambda A + \mu B$.

* Kronecker, Leopold (1823-1891). German algebraist and algebraic number theorist. Capelli, A. (1855-1910). Italian mathematician.

The numbers a_{kl} called the *coefficients of system* (1) are specified. We shall also say that system (1) is defined by a matrix of its coefficients:

$$A = \| a_{kl} \| = \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{vmatrix}. \quad (2)$$

We shall be interested in the problem of solvability of system (1) for each vector (system of numbers)

$$\mathbf{y} = (y_1, \dots, y_n).$$

A system of numbers (vector)

$$\mathbf{x} = (x_1, \dots, x_n)$$

is called the *solution of system of equations* (1) if the numbers x_j satisfy these equations.

Theorem 1. *If the determinant of system (1) is not equal to zero:*

$$\Delta = | a_{kl} | \neq 0,$$

*then system (1) has a unique solution for any vector $\mathbf{y}$ which is computed by Cramer's formulas**

$$x_j = \Delta^j / \Delta \quad (j = 1, \dots, n), \quad (3)$$

where Δ^j is a determinant obtained from the determinant Δ by replacing in it the numbers of the j th column by the respective numbers $y_1, \dots, y_n$:

$$\Delta^j = \begin{vmatrix} a_{11} & \dots & a_{1, j-1} & y_1 & a_{1, j+1} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & \dots & a_{n, j-1} & y_n & a_{n, j+1} & \dots & a_{nn} \end{vmatrix}. \quad (4)$$

Hence,

$$x_j = \frac{1}{\Delta} \sum_{s=1}^n A_{sj} y_s \quad (j = 1, \dots, n), \quad (3')$$

where A_{sj} is the cofactor of the element a_{sj} in the determinant Δ .

* Cramer, Gabriel (1704-1752). Swiss mathematician and physicist.

is called *homogeneous*. It is a particular case of system (1) for $y_1 = \dots = y_n = 0$. It is obvious that the zero vector

$$x_1 = 0, \dots, x_n = 0$$

satisfies homogeneous system (5). But it may happen that homogeneous system (5) is satisfied by a nonzero vector $\mathbf{x} = (x_1, \dots, x_n)$, i.e. by a vector having at least one component $x_i \neq 0$. This vector is called a *nontrivial solution of homogeneous system (5)*, and zero vector is therefore called a *trivial solution of homogeneous system (5)*.

Theorem 2. *If the determinant Δ of homogeneous system (5) is not zero ($\Delta \neq 0$), then this system has only a trivial solution.*

Indeed, by property (d), all the determinants $\Delta^j = 0$ (see (4)), therefore, by virtue of equalities (3), $x_j = 0$ ($j = 1, \dots, n$).

Theorem 3. *If system of equations (5) has a nontrivial solution, then its determinant Δ is necessarily equal to zero ($\Delta = 0$).*

Indeed, if Δ were not equal to zero, then, by Theorem 2, system (5) would have only a trivial solution.

In the foregoing paragraphs we investigated linear system (1) for the case when its determinant $\Delta \neq 0$. In this case it was shown (Theorem 1) that system (1) has for any right-hand member of $\mathbf{y} = (y_1, \dots, y_n)$ a unique solution which is computed by formulas (3).

We now pass over to the case when the determinant of this system is zero ($\Delta = 0$). We shall assume that at least one element of matrix A (see (2)) is nonzero and denote rank A by k ($k = \text{rank } A$). Thus, $1 \leq k < n$.

Our purpose is to prove the following rules (they were formulated and proved explicitly by Kronecker and Capelli).

If we want to solve system (1) for which it is known that the rank of the matrix A of its coefficients is equal to k , we must find out the rank of the *augmented matrix*

$$B = \left\| \begin{array}{cccc} a_{11} & \dots & a_{1n} & y_1 \\ \dots & \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} & y_n \end{array} \right\|,$$

obtained by adding to A the column

$$\left\| \begin{array}{c} y_1 \\ y_2 \\ \dots \\ y_n \end{array} \right\| = \left(\begin{array}{c} y_1 \\ y_2 \\ \dots \\ y_n \end{array} \right).$$

(1) If rank B is greater than rank A (rank $B > \text{rank } A = k$), then system (1) has no solution at all. It is *contradictory*, since there is no vector $\mathbf{x} = (x_1, \dots, x_n)$, satisfying simultaneously all equations (1).

(2) If rank B equals rank A (rank $B = \text{rank } A = k$), then system (1) has solutions. In order to find them, we must choose from system (1) some k equations the matrix of coefficients of which has a rank k , and solve these k equations. This system of k equations will have an infinite set of solutions, but they can be written in a visible manner.

In this case any solution of the k chosen equations is automatically a solution of the remaining $n - k$ equations of system (1).

Rules (1) and (2) exhaust all possible situations, since rank B cannot be less than k .

As a matter of fact, matrix A , by hypothesis, generates a nonzero determinant of order k which is also generated by matrix B .

Example 1. The system

$$\left. \begin{array}{l} x + y = 1, \\ x - y = 2 \end{array} \right\}$$

has a determinant

$$\Delta = \begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix} = -2 \neq 0$$

and therefore has a unique solution which can be computed by the formulas

$$x = \Delta^1 / \Delta, \quad y = \Delta^2 / \Delta,$$

where

$$\Delta^1 = \begin{vmatrix} 1 & 1 \\ 2 & -1 \end{vmatrix} = -3, \quad \Delta^2 = \begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix} = 1,$$

i.e.

$$x = \frac{-3}{-2} = \frac{3}{2}, \quad y = \frac{1}{-2} = -\frac{1}{2}.$$

Example 2. The system

$$\left. \begin{aligned} x + y &= 1, \\ x + y &= 2 \end{aligned} \right\} \quad (6)$$

has a determinant equal to zero. The matrix

$$A = \begin{vmatrix} 1 & 1 \\ 1 & 1 \end{vmatrix}$$

has a rank $A = 1$. And the matrix

$$B = \begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \end{vmatrix}$$

has a rank $B = 2$. Since $\text{rank } B > \text{rank } A$, system (6) has no solution. However, this is seen without our theory: one and the same number cannot be simultaneously equal to 1 and to 2.

Example 3. The system

$$\left. \begin{aligned} 2x + 2y &= 2, \\ 3x + 3y &= 3 \end{aligned} \right\} \quad (7)$$

has a determinant $\Delta = 0$. The matrix

$$A = \begin{vmatrix} 2 & 2 \\ 3 & 3 \end{vmatrix}$$

has a rank $A = 1$. The augmented matrix

$$B = \begin{vmatrix} 2 & 2 & 2 \\ 3 & 3 & 3 \end{vmatrix}$$

also has a rank (B) equal to 1. Since $\text{rank } A = \text{rank } B = 1$, we take one equation

$$2x + 2y = 2. \quad (8)$$

As is seen, the coefficient of y is not equal to zero, therefore this equation can be solved with respect to y :

$$y = \frac{2-2x}{2} = 1-x. \quad (9)$$

Formula (9) yields all solutions of equation (8). We can assign any value to x ($-\infty < x < \infty$) and compute the corresponding value of y by formula (9), thus obtaining a system (vector) (x, y) satisfying equation (8). The set of all systems $(x, 1 - x)$, where $x \in (-\infty, \infty)$, forms the set of all solutions of equation (8). These solutions are automatically solutions of the second equation of system (7), since $\text{rank } A = \text{rank } B$. In this case the result is obvious without applying the theory of matrix ranks. The coefficients of equations (7) as well as the right-hand members of these equations are respectively proportional, therefore it is clear that any solution of one of these equations is at the same time a solution of the other.

Example 4. The system

$$\left. \begin{aligned} x + y + z &= 1, \\ 2x + y + 2z &= 1, \\ x + y + 3z &= 2 \end{aligned} \right\}$$

has a determinant

$$\Delta = \begin{vmatrix} 1 & 1 & 1 \\ 2 & 1 & 2 \\ 1 & 1 & 3 \end{vmatrix} = -2 \neq 0$$

and, therefore, has a unique solution which can be computed by the formulas

$$\begin{aligned} x &= \frac{\Delta^1}{\Delta} = \frac{-1}{-2} \begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \\ 2 & 1 & 3 \end{vmatrix} = -\frac{1}{2}, & y &= \frac{\Delta^2}{\Delta} \\ &= \frac{-1}{-2} \begin{vmatrix} 1 & 1 & 1 \\ 2 & 1 & 2 \\ 1 & 2 & 3 \end{vmatrix} = 1, & z &= \frac{\Delta^3}{\Delta} = \frac{-1}{-2} \begin{vmatrix} 1 & 1 & 1 \\ 2 & 1 & 1 \\ 1 & 1 & 2 \end{vmatrix} = \frac{1}{2}. \end{aligned}$$

Example 5. The system

$$\left. \begin{aligned} x + y + z &= 1, \\ x + y + 2z &= 1, \\ x + y + 3z &= 2 \end{aligned} \right\}$$

has a determinant

$$\Delta = \begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \\ 1 & 1 & 3 \end{vmatrix} = 0.$$

The matrix

$$A = \begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \\ 1 & 1 & 3 \end{vmatrix}$$

has a rank $A = 2$, since $\Delta = 0$, but there exists a nonzero second-order determinant generated by matrix A . For instance,

$$\begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix} = 1 \neq 0.$$

The matrix

$$B = \begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 2 & 1 \\ 1 & 1 & 3 & 2 \end{vmatrix}$$

has a rank $B = 3$, since the determinant generated by this matrix

$$\begin{vmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 3 & 2 \end{vmatrix} = 1 \neq 0.$$

Since rank $B > \text{rank } A$, the system has no solution.

Example 6. The system

$$\left. \begin{aligned} x + y + z &= 1, \\ x + y + 2z &= 1, \\ 2x + 2y + 4z &= 2 \end{aligned} \right\} \quad (10)$$

has a determinant

$$\Delta = \begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \\ 2 & 2 & 4 \end{vmatrix} = 0.$$

It is easy to compute that the matrices

$$A = \begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \\ 2 & 2 & 4 \end{vmatrix}, \quad B = \begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 2 & 1 \\ 2 & 2 & 4 & 2 \end{vmatrix}$$

have equal ranks, where $\text{rank } A = \text{rank } B = 2$. Let us choose from system (10) two equations so that the rank of the matrix A' formed from the coefficients of these equations is equal to 2. In this case it is possible to take the first and second equations, or the first and the third. Thus, let us consider the system

$$\begin{cases} x + y + z = 1, \\ x + y + 2z = 1. \end{cases} \quad (11)$$

We transpose to the right-hand sides of these equations one of the unknowns so that the coefficients of the remaining unknowns form a matrix A'' whose rank $A'' = 2$. Here we may transpose either x or y .

Hence, the nonhomogeneous system

$$\begin{cases} x + z = 1 - y, \\ x + 2z = 1 - y \end{cases} \quad (12)$$

has a determinant

$$\Delta'' = \begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix} = 1 \neq 0,$$

therefore it has a unique solution for any right-hand side:

$$x = \frac{1}{\Delta''} \begin{vmatrix} 1 - y & 1 \\ 1 - y & 2 \end{vmatrix} = 1 - y, \quad z = \frac{1}{\Delta''} \begin{vmatrix} 1 & 1 - y \\ 1 & 1 - y \end{vmatrix} = 0.$$

Thus, for any y the sets of three numbers $(1 - y, y, 0)$ yield all solutions of system (12) and, automatically, the solutions of the third equation of system (10) (this equation is obtained from the second by multiplying it by 2).

Theorem 4. *If system (1) is compatible, i.e. if it has at least one solution $\mathbf{x}$, then it is necessary that rank B be equal to rank A .*

Let us set up a system with the unknowns $z_1, z_2, \dots, z_{k+1}$:

$$\left. \begin{aligned} a_{11}z_1 + \dots + a_{1k}z_k + \lambda_1 z_{k+1} &= 0, \\ \dots & \\ a_{k+1,1}z_1 + \dots + a_{k+1,k}z_k + \lambda_{k+1}z_{k+1} &= 0. \end{aligned} \right\} \quad (15)$$

On the basis of (13) and (14), this system is satisfied by the numbers $x_1, \dots, x_k, 1$ one of which, at all events, is not equal to zero. But then the determinant of homogeneous system (15) is zero (see Theorem 3), i.e.

$$\begin{aligned} 0 &= \begin{vmatrix} a_{11} & \dots & a_{1k} & a_{1,k+1}x_{k+1} + \dots + a_{1n}x_n & -y_1 \\ \dots & \dots & \dots & \dots & \dots \\ a_{k+1,1} & \dots & a_{k+1,k} & a_{k+1,k+1}x_{k+1} + \dots + a_{k+1,n}x_n & -y_{k+1} \end{vmatrix} \\ &= \sum_{s=k+1}^n x_s \begin{vmatrix} a_{11} & \dots & a_{1k} & a_{1s} \\ \dots & \dots & \dots & \dots \\ a_{k+1,1} & \dots & a_{k+1,k} & a_{k+1,s} \end{vmatrix} \\ &\quad - \begin{vmatrix} a_{11} & \dots & a_{1k} & y_1 \\ \dots & \dots & \dots & \dots \\ a_{k+1,1} & \dots & a_{k+1,k} & y_{k+1} \end{vmatrix} \\ &= - \begin{vmatrix} a_{11} & \dots & a_{1k} & y_1 \\ \dots & \dots & \dots & \dots \\ a_{k+1,1} & \dots & a_{k+1,k} & y_{k+1} \end{vmatrix}, \quad (16) \end{aligned}$$

since the determinants (of order $k+1$) entering into the sum $\sum_s$ are equal to zero, as a matter of fact, the rank of matrix A is equal to k .

Thus, we have proved that any determinant of order $k+1$ generated by matrix B is equal to zero, which was to be proved.

Now we are going to justify rule (2). Since the rank of matrix A is equal to k ($\text{rank } A = k$), system (1) contains k equations, the matrix of coefficients of which generates a nonzero determinant of order k . By renumbering the equations and unknowns, it is possible to achieve

that $\text{rank } A = \text{rank } B = k$. For the sake of definiteness, let us prove that it is the solution of the $(k + 1)$ th equation. Consider the first $(k + 1)$ equations of system (1) written in the form of (13). We have to prove that any solution $\mathbf{x}$ of the first k equations of (13) is automatically the solution of $(k + 1)$ th equation in (13). Let $\mathbf{x}$ be a vector satisfying the first k equations in (13). We set up equations (15) with respect to the unknowns $z_1, \dots, z_{k+1}$, where the numbers $\lambda_1, \dots, \lambda_{k+1}$ are computed by formulas (14) in terms of the components $x_{k+1}, \dots, x_n$ of the vector $\mathbf{x}$. The determinant of system (15) is equal to zero. This is seen from equations (16) which should be read from right to left. By hypothesis, the determinant on the right is zero. But then system (15) has a nontrivial solution $z_1, z_2, \dots, z_{k+1}$. Here the number $z_{k+1} \neq 0$ because, if we assume that system (15) has a solution of the form $z_1, \dots, z_k, 0$, then the numbers $z_1, \dots, z_k$ must vanish, since the determinant $\sigma \neq 0$. But then $z_1 = \dots = z_k = z_{k+1} = 0$ and the system $z_1, \dots, z_{k+1}$ would be a trivial one. Since system (15) is a homogeneous one, not only the numbers $z_1, \dots, z_{k+1}$, but also the numbers

$$z'_1 = z_1/z_{k+1}, \dots, z'_k = z_k/z_{k+1}, 1$$

possess the same property ($z_{k+1} \neq 0$). But then the numbers $z'_1, \dots, z'_k$ satisfy the system of the first k equations of (13), which has a nonzero determinant ($\sigma \neq 0$). We already know that this system has solutions $x_1, \dots, x_k$, and, by virtue of uniqueness,

$$z'_1 = x_1, \dots, z'_k = x_k.$$

Considering the last equation of (15), we see that it is satisfied by the numbers $(x_1, \dots, x_k, 1)$, i.e. the numbers $(x_1, \dots, x_k)$ satisfy the $(k + 1)$ th equation of system (13) and, by virtue of (14), the vector under consideration $\mathbf{x} = (x_1, \dots, x_k, x_{k+1}, \dots, x_n)$ satisfies the $(k + 1)$ th equation of system (1). This proves statement (2).

Remark 1. Hereby we recommend the following method for solving systems of linear equations, which is, as a matter of fact, the *method of elimination of unknowns*.

Form matrix B :

$$B = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 1 & 2 & 3 & 1 \\ 1 & 0 & 3 & 4 & 2 \\ 1 & 1 & 5 & 6 & 1 \end{pmatrix},$$

in which, as we see, the last column consists of right-hand members of the given system. Multiplying the first row by (-1) and adding it to the third and fourth rows, we get the matrix

$$B_1' = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 1 & 2 & 3 & 1 \\ 0 & -2 & 0 & 0 & -3 \\ 0 & -1 & 2 & 2 & -4 \end{pmatrix}.$$

In this matrix the elements of the third row (excepting one) which are coefficients of the unknowns are equal to zero. Let us interchange the second and third rows. Then the nonzero element of the third row will be found on the principal diagonal:

$$B_1'' = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & -2 & 0 & 0 & -3 \\ 0 & 1 & 2 & 3 & 1 \\ 0 & -1 & 2 & 2 & -4 \end{pmatrix}.$$

In order to simplify the notation, we multiply the second row by -1 :

$$B_1 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 2 & 0 & 0 & 3 \\ 0 & 1 & 2 & 3 & 1 \\ 0 & -1 & 2 & 2 & -4 \end{pmatrix}.$$

Further transformations of the matrices are obvious:

$$\begin{aligned}
 B_1 \Rightarrow B_2 &= \begin{pmatrix} 1 & 0 & 3 & 4 & 2 \\ 0 & 2 & 0 & 0 & 3 \\ 0 & 0 & 2 & 3 & -1/2 \\ 0 & 0 & 2 & 2 & -5/2 \end{pmatrix} \Rightarrow B_3 \\
 &= \begin{pmatrix} 1 & 0 & 3 & 4 & 2 \\ 0 & 2 & 0 & 0 & 3 \\ 0 & 0 & 2 & 3 & -1/2 \\ 0 & 0 & 0 & -1 & -2 \end{pmatrix} \Rightarrow B_4 = \begin{pmatrix} 1 & 0 & 3 & 4 & 2 \\ 0 & 2 & 0 & 0 & 3 \\ 0 & 0 & 2 & 3 & -1/2 \\ 0 & 0 & 0 & 1 & 2 \end{pmatrix} \Rightarrow B_5 \\
 &= \begin{pmatrix} 1 & 0 & 3 & 0 & -6 \\ 0 & 2 & 0 & 0 & 3 \\ 0 & 0 & 2 & 0 & -13/2 \\ 0 & 0 & 0 & 1 & 2 \end{pmatrix} \Rightarrow B_6 = \begin{pmatrix} 1 & 0 & 0 & 0 & 15/4 \\ 0 & 2 & 0 & 0 & 3 \\ 0 & 0 & 2 & 0 & -13/2 \\ 0 & 0 & 0 & 1 & 2 \end{pmatrix}.
 \end{aligned}$$

Hence, $x_4 = 2$, $x_3 = -13/4$, $x_2 = 3/2$, $x_1 = 15/4$. In order not to commit an error, it is recommended to effect a check by substituting the obtained values into the initial equations of the given system.

Consider Example 5 from this point of view:

$$\begin{aligned}
 &\left. \begin{aligned} x_1 + x_2 + x_3 &= 1, \\ x_1 + x_2 + 2x_3 &= 1, \\ x_1 + x_2 + 3x_3 &= 2, \end{aligned} \right\} \\
 B = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 2 & 1 \\ 1 & 1 & 3 & 2 \end{pmatrix} &\Rightarrow B_1 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 2 & 1 \end{pmatrix} \Rightarrow B_2 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.
 \end{aligned}$$

Hence, the initial system is equivalent to the following one:

$$\left. \begin{aligned} x_1 + x_2 + x_3 &= 1, \\ 0 \cdot x_1 + 0 \cdot x_2 + x_3 &= 0, \\ 0 \cdot x_1 + 0 \cdot x_2 + 0 \cdot x_3 &= 1. \end{aligned} \right\}$$

The constant term of the last row is equal to unity, and the coefficients of the unknowns are zero, therefore the system is incompatible.

Finally, in Example 6

$$\begin{aligned}
 B = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 2 & 1 \\ 2 & 2 & 4 & 2 \end{pmatrix} &\Rightarrow B_1 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 2 & 0 \end{pmatrix} \Rightarrow B_2 \\
 &= \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \Rightarrow B_3 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \Rightarrow B_4 = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.
 \end{aligned}$$

Hence, $x_3 = 0$, $x_1 + x_2 = 1$, i.e. the system has an infinite set of solutions: $x_1 = x_1$, $x_2 = 1 - x_1$, $x_3 = 0$, where x_1 is any number ($-\infty < x_1 < \infty$).

Remark 2. And if we are interested only in the rank of matrix B , then the above indicated operations $B \Rightarrow B'$ should be extended not only to the rows, but also to the columns of the matrix. If in the process of these transformations there appears a row or a column consisting only of zeros, then they should be deleted and a matrix of a smaller dimension should be given further consideration.

This method, known as *Gauss method** is illustrated by the following examples.

Example 8. Find the rank of the matrix

$$B = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 1 & 2 & 1 & 2 & 1 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 4 & 5 \end{pmatrix}.$$

It is clear that the rank of matrix B is not greater than 4. In this case $a_{11} = 1 \neq 0$. Multiplying the first

* Gauss, Carl Friedrich (1777-1855). German mathematician, usually considered along with Aristotle and Newton to be one of the three greatest mathematicians of all time. Made important contributions to algebra, analysis, geometry, number theory, numerical analysis, probability, and statistics, as well as astronomy and physics.

row by (-1) and adding it to the third row, we get

$$B_1 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & -2 & -2 & -4 & -4 & -6 \\ 0 & 0 & 0 & 0 & 0 & 4 & 5 \end{pmatrix}.$$

Multiplying now the first column by the corresponding numbers and adding it to the remaining columns, we obtain the matrix

$$B_2 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & -2 & -2 & -4 & -4 & -6 \\ 0 & 0 & 0 & 0 & 0 & 4 & 5 \end{pmatrix}.$$

The second column now consists of zeros, except for the element $a_{22} = 1 \neq 0$. Multiplying the second column by (-1) and adding it to the fourth, sixth, and seventh columns, we get

$$\begin{aligned} B_3 &= \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -2 & -2 & -4 & -4 & -6 \\ 0 & 0 & 0 & 0 & 0 & 4 & 5 \end{pmatrix} \Rightarrow B_4 \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 4 & 5 \end{pmatrix} \Rightarrow B_5 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -2 & 0 & 0 \\ 0 & 0 & 0 & 4 & 5 \end{pmatrix} \Rightarrow B_6 \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -2 & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 \end{pmatrix} \Rightarrow B_7 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}. \end{aligned}$$

The fourth-order determinant of matrix B_7 is not equal to zero, consequently,

$$\text{rank } B = \text{rank } B_7 = 4.$$

Example 9. Find the rank of the matrix

$$\begin{aligned}
 B &= \begin{pmatrix} 1 & 2 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \end{pmatrix} \Rightarrow B_1 = \begin{pmatrix} 1 & 2 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & -1 & -1 & -1 \end{pmatrix} \Rightarrow B_2 \\
 &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & -1 & -1 & -1 \end{pmatrix} \Rightarrow B_3 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \Rightarrow B_4 \\
 &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \end{pmatrix} \Rightarrow B_5 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \Rightarrow B_6 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},
 \end{aligned}$$

i.e. the rank of matrix B equals two.

Sec. 5. Three-dimensional Space. Vectors. Cartesian System of Coordinates *

In this section we shall consider a real space. The concept of the vector in a real space is familiar to the reader from elementary geometry.

The *vector* (in a real space) is defined as a directed line-segment $\overrightarrow{AB}$ with initial point A and terminal point B that can be displaced parallel to itself. Thus, it is considered that two directed vectors $\overrightarrow{AB}$ and $\overrightarrow{A_1B_1}$ having equal lengths ($|AB| = |A_1B_1|$) and one and the same direction define one and the same vector $\mathbf{a}$, and in this sense we write $\mathbf{a} = \overrightarrow{AB} = \overrightarrow{A_1B_1}$ (Fig. 2).

The *length* $|\overrightarrow{AB}|$ of the vector $\overrightarrow{AB}$ is a nonnegative number equal to the length of the line-segment AB joining the points A and B . We shall also write $|\overrightarrow{AB}| = |AB|$.

Vectors lying on parallel straight lines or on one and the same straight line are termed *collinear*.

* Note that the present book first deals with the scalar product of two vectors, then with analytical geometry of the straight line and plane, and after this, in Secs. 11-13, the concepts of the vector and scalar triple products of vectors are set forth. If desired, these sections may immediately follow Sec. 6.

If the points A and B coincide, then $\overrightarrow{AB} = \overrightarrow{AA} = \mathbf{0}$ is also regarded as a vector, the *null vector*. Its length is equal to zero ($|\mathbf{0}| = 0$), and direction has no sense for it. However, for the sake of generality of the rules

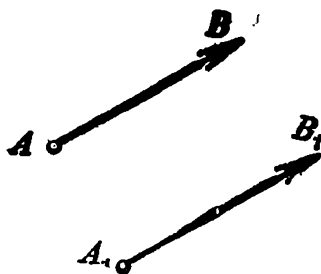


Fig. 2.

of vector algebra, it is agreed that a pair of coincident points is to be regarded as a null vector. The latter is considered collinear with any vector.

The projection of a point A on a straight line L (Fig. 3) is a point A' at which the line L cuts the plane perpendicular to L and passing through the point A .

Let there be given a directed straight line L (Fig. 4) and vector $\mathbf{a} = \overrightarrow{AB}$. The projection of the vector $\mathbf{a} = \overrightarrow{AB}$ on the directed line L is the vector $\overrightarrow{A'B'}$, where A' , B' are the projections of the points A , B on L , respectively.

With the given directed line L , the projections $\overrightarrow{A'B'}$ of any vectors $\overrightarrow{AB}$ lie in L and are either in the same direction as L , or in the opposite direction. This enables us to express the projections of vectors on a given directed line also by numbers and, hence, we can give another definition, this time an algebraic one.

The projection of the vector $\mathbf{a} = \overrightarrow{AB}$ on the directed straight line L is defined as the product of the length of the vector $\mathbf{a} = \overrightarrow{AB}$ by the cosine of the angle ω between

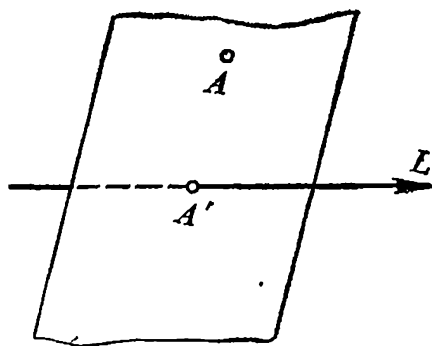


Fig. 3.

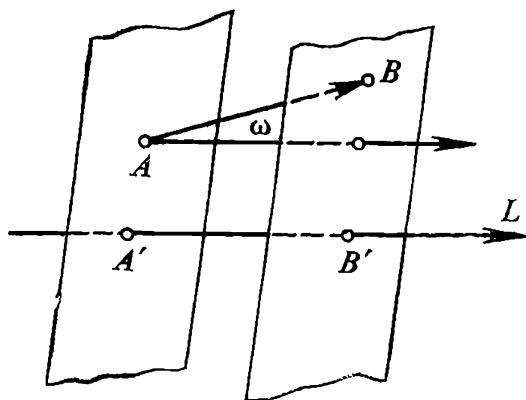


Fig. 4.

the vector $\mathbf{a}$ and the direction of L :

$$\text{pr}_L \mathbf{a} = |\mathbf{a}| \cos (\mathbf{a}, L) = |\mathbf{a}| \cos \omega \quad (0 \leq \omega \leq \pi). \quad (1)$$

The transition from the first definition to the second is carried out in the following way: $\text{pr}_L \mathbf{a}$ is a number equal to the length $|A'B'|$ taken with the sign $+$ or $-$ depending on whether the vector $\overrightarrow{A'B'}$ is in the same direction as L or in the opposite direction.

In geometry, we consider addition and subtraction of vectors, as well as their multiplication by real numbers. By definition, the product $\alpha \mathbf{a} = \alpha \mathbf{a}$ of a vector $\mathbf{a}$ by a number α or a number α by a vector $\mathbf{a}$ is a vector whose length $|\alpha \mathbf{a}| = |\alpha| \cdot |\mathbf{a}|$, and the direction coincides

with that of $\mathbf{a}$, if $\alpha > 0$ or opposite to $\mathbf{a}$ if $\alpha < 0$. For $\alpha = 0$ the length $|\alpha \mathbf{a}|$ is zero and the vector $\alpha \mathbf{a}$ turns into a null vector (point) having no direction.

By definition, a finite number of vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, $\dots$ are added according to the rule for closing the "chain"

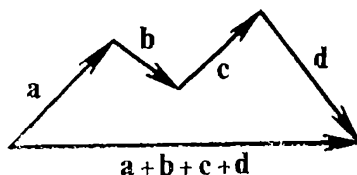


Fig. 5.

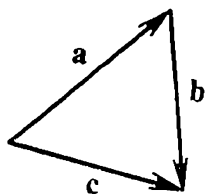


Fig. 6.

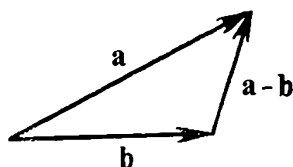


Fig. 7.

of these vectors. Figure 5 and 6 recall the way this is done. Figure 7 shows how vectors are subtracted.

The projections of vectors $\mathbf{a}$, $\mathbf{b}$ on a given direction L possess the following properties:

$$\text{pr}_L \mathbf{a} + \text{pr}_L \mathbf{b} = \text{pr}_L (\mathbf{a} + \mathbf{b}), \quad (2)$$

$$\text{pr}_L (\alpha \mathbf{a}) = \alpha \text{pr}_L \mathbf{a}. \quad (3)$$

Property (2) is illustrated in Fig. 8:

$$\text{pr}_L \mathbf{a} + \text{pr}_L \mathbf{b} = \overrightarrow{A'B'} + \overrightarrow{B'C'} = \overrightarrow{A'C'} = \text{pr}_L \mathbf{c} = \text{pr}_L (\mathbf{a} + \mathbf{b}).$$

Since $\mathbf{a} = (\mathbf{a} - \mathbf{b}) + \mathbf{b}$ (see Fig. 7), we have

$$\text{pr}_L (\mathbf{a} - \mathbf{b}) + \text{pr}_L \mathbf{b} = \text{pr}_L \mathbf{a},$$

and, consequently,

$$\text{pr}_L \mathbf{a} - \text{pr}_L \mathbf{b} = \text{pr}_L (\mathbf{a} - \mathbf{b}). \quad (2')$$

Let us prove property (3). Taking the angle between the vector $\mathbf{a}$ and the direction L equal to ω , we have

for $\alpha > 0$: $\text{pr}_L(\alpha\mathbf{a}) = |\alpha\mathbf{a}| \cos \omega = \alpha |\mathbf{a}| \cos \omega = \alpha \text{pr}_L \mathbf{a}$;

for $\alpha < 0$: $\text{pr}_L(\alpha\mathbf{a}) = |\alpha\mathbf{a}| \cos(\pi - \omega)$
 $= -\alpha |\mathbf{a}| \cos(\pi - \omega) = \alpha |\mathbf{a}| \cos \omega = \alpha \text{pr}_L \mathbf{a}$.

As is known, for $\alpha < 0$ the vector $\alpha\mathbf{a}$ is in the direction

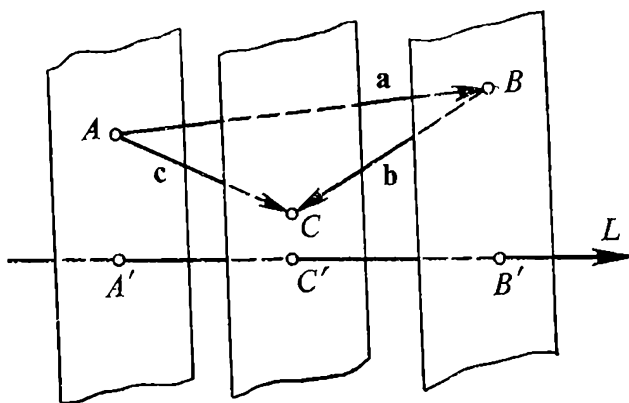


Fig. 8.

opposite to that of $\mathbf{a}$, and if $\mathbf{a}$ forms an angle ω with L , then $\alpha\mathbf{a}$ is at an angle $\pi - \omega$ to L .

For $\alpha = 0$ both sides of (3) vanish.

A *scalar product of two vectors* $\mathbf{a}$ and $\mathbf{b}$ is defined as a number $(\mathbf{a}, \mathbf{b})$ equal to the product of the lengths of these vectors multiplied by the cosine of the angle ω between them:

$$\mathbf{ab} = (\mathbf{a}, \mathbf{b}) = |\mathbf{a}| |\mathbf{b}| \cos(\mathbf{a}, \mathbf{b}) = |\mathbf{a}| |\mathbf{b}| \cos \omega. \quad (4)$$

Obviously, we may also say that the *scalar product of vectors* $\mathbf{a}$ and $\mathbf{b}$ is the product of the length of the vector $|\mathbf{b}|$ by the projection of the vector $\mathbf{a}$ on the direction $\mathbf{b}$ or the product of the length $|\mathbf{a}|$ by the projection $\mathbf{b}$ on the direction $\mathbf{a}$:

$$\mathbf{ab} = (\mathbf{a}, \mathbf{b}) = |\mathbf{a}| \text{pr}_{\mathbf{a}} \mathbf{b} = |\mathbf{b}| \text{pr}_{\mathbf{b}} \mathbf{a}.$$

The scalar product possesses the following properties:

$$(\mathbf{a}, \mathbf{b}) = (\mathbf{b}, \mathbf{a}), \quad (5)$$

$$(\mathbf{a}, \mathbf{b} + \mathbf{c}) = (\mathbf{a}, \mathbf{b}) + (\mathbf{a}, \mathbf{c}), \quad (6)$$

$$(\mathbf{a}, \alpha \mathbf{b}) = \alpha (\mathbf{a}, \mathbf{b}). \quad (7)$$

Equation (5) follows directly from the definition of the scalar product.

Equation (6) is proved as follows:

$$\begin{aligned} (\mathbf{a}, \mathbf{b} + \mathbf{c}) &= |\mathbf{a}| \operatorname{pr}_{\mathbf{a}} (\mathbf{b} + \mathbf{c}) \\ &= |\mathbf{a}| \operatorname{pr}_{\mathbf{a}} \mathbf{b} + |\mathbf{a}| \operatorname{pr}_{\mathbf{a}} \mathbf{c} = (\mathbf{a}, \mathbf{b}) + (\mathbf{a}, \mathbf{c}). \end{aligned}$$

And this is how equation (7) is proved:

$$(\mathbf{a}, \alpha \mathbf{b}) = |\mathbf{a}| \operatorname{pr}_{\mathbf{a}} (\alpha \mathbf{b}) = |\mathbf{a}| \alpha \operatorname{pr}_{\mathbf{a}} \mathbf{b} = \alpha (\mathbf{a}, \mathbf{b}).$$

Taking into consideration (5), it follows from (6) and (7):

$$(\mathbf{a} + \mathbf{b}, \mathbf{c}) = (\mathbf{a}, \mathbf{c}) + (\mathbf{b}, \mathbf{c}), \quad (6')$$

$$(\alpha \mathbf{a}, \mathbf{c}) = \alpha (\mathbf{a}, \mathbf{c}). \quad (7')$$

Now we pass over to an analytical description of vectors and points in space, i.e. to describing them with the aid of numbers. Let us introduce in space a *rectangular coordinate system* x, y, z , i.e. three mutually perpendicular directed lines passing through a certain point O called *coordinate axes* x, y, z (Fig. 9). It is assumed that for this coordinate system a unit line-segment is chosen by means of which all other segments are measured. The point O is called the *origin* (of coordinates).

Let us specify an arbitrary point A of a three-dimensional space. The directed line-segment $\overrightarrow{OA}$ is called the *radius vector* of the point A . The radius vector, in its turn, defines the vector $\mathbf{a}$ ($\mathbf{a} = \overrightarrow{OA}$), which can be translated in space. We denote the projections of the radius vector $\mathbf{a}$ on the axes x, y, z by x, y, z , respectively. These are *coordinates* of the point A , which are called the *abscissa* (x), *ordinate* (y), and *z-coordinate* of the point A .

Between points A in space and their radii vectors $\overrightarrow{OA}$ or, which is the same, triples of numbers (x, y, z) which

are the coordinates of the point A or the projections of $\overrightarrow{OA}$ on the coordinate axes there is one-to-one correspondence. That is why there will be no confusion if we call the

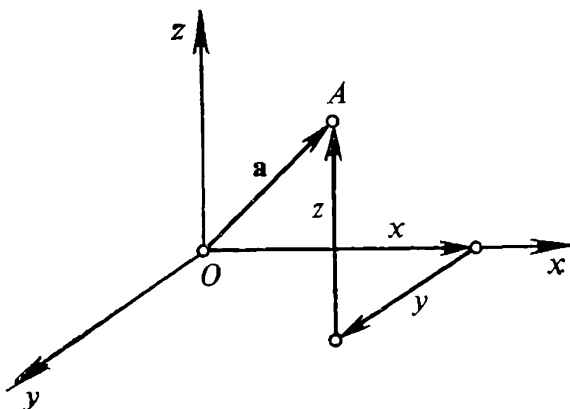


Fig. 9.

triple of numbers (x, y, z) either a point A having these numbers as its coordinates, or a vector $\mathbf{a}$ having these numbers as its projections.

From the definition of the vector as a directed line-segment which can be translated in space, it follows that two vectors $\mathbf{a} = (x_1, y_1, z_1)$ and $\mathbf{b} = (x_2, y_2, z_2)$ are equal if, and only if, the equalities $x_1 = x_2$, $y_1 = y_2$, $z_1 = z_2$ are fulfilled simultaneously. The following equations hold true:

$$(x, y, z) \pm (x', y', z') = (x \pm x', y \pm y', z \pm z'), \quad (8)$$

$$\alpha (x, y, z) = (\alpha x, \alpha y, \alpha z). \quad (9)$$

Equation (8) follows from that the projection of a sum or a difference of vectors (on one of the coordinate axes) is equal to the sum or difference of the projections of the terms.

Equation (9) follows from that the projection of the vector $\alpha \mathbf{a}$ (on one of the coordinate axes) is equal to the product of α by the projection of $\mathbf{a}$.

Let us denote by $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ unit vectors (of length equal to 1) having the same direction as the axes x , y , z ,

respectively. The vectors $\mathbf{i}, \mathbf{j}, \mathbf{k}$ are called *basis vectors* of the axes x, y, z . An arbitrary vector (x, y, z) can be written in the form

$$(x, y, z) = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}. \quad (10)$$

Indeed,

$$\mathbf{i} = (1, 0, 0), \quad \mathbf{j} = (0, 1, 0), \quad \mathbf{k} = (0, 0, 1).$$

Therefore, by (8) and (9),

$$\begin{aligned} x\mathbf{i} + y\mathbf{j} + z\mathbf{k} &= x(1, 0, 0) + y(0, 1, 0) + z(0, 0, 1) \\ &= (x, 0, 0) + (0, y, 0) + (0, 0, z) = (x, y, z). \end{aligned}$$

Let us note the equations relating the scalar products of the unit vectors of the coordinate axes:

$$\mathbf{i}\mathbf{i} = \mathbf{j}\mathbf{j} = \mathbf{k}\mathbf{k} = 1, \quad \mathbf{i}\mathbf{j} = \mathbf{i}\mathbf{k} = \mathbf{j}\mathbf{k} = 0.$$

Let now $\mathbf{a} = (x, y, z)$, $\mathbf{b} = (x', y', z')$. Then

$$\mathbf{a}\mathbf{b} = (\mathbf{a}, \mathbf{b}) = xx' + yy' + zz'. \quad (11)$$

Indeed, by virtue of (6), (7), (6'), (7'),

$$\begin{aligned} \mathbf{a}\mathbf{b} &= (x\mathbf{i} + y\mathbf{j} + z\mathbf{k})(x'\mathbf{i} + y'\mathbf{j} + z'\mathbf{k}) \\ &= xx'\mathbf{i}\mathbf{i} + xy'\mathbf{i}\mathbf{j} + xz'\mathbf{i}\mathbf{k} + yx'\mathbf{j}\mathbf{i} + yy'\mathbf{j}\mathbf{j} + yz'\mathbf{j}\mathbf{k} \\ &\quad + zx'\mathbf{k}\mathbf{i} + zy'\mathbf{k}\mathbf{j} + zz'\mathbf{k}\mathbf{k} = xx' + yy' + zz' \end{aligned}$$

In particular, making $\mathbf{b} = \mathbf{a}$ in this formula, we get

$$|\mathbf{a}|^2 = \mathbf{a}\mathbf{a} = x^2 + y^2 + z^2,$$

whence the length of the vector $\mathbf{a}$ is equal to

$$|\mathbf{a}| = +\sqrt{x^2 + y^2 + z^2}.$$

Hence, the distance between the points $\mathbf{a} = (x, y, z)$ and $\mathbf{b} = (x', y', z')$ is equal to (Fig. 10)

$$|\mathbf{AB}| = |\mathbf{b} - \mathbf{a}| = +\sqrt{(x' - x)^2 + (y' - y)^2 + (z' - z)^2}.$$

For further treatment of the material it is important to sum up what has been said. To this effect, let us introduce somewhat different designation of coordinates. Namely, we are going to introduce in space a rectangular system of coordinates x_1, x_2, x_3 . Then every point of

space will be represented by a triple of numbers

$$\mathbf{x} = (x_1, x_2, x_3).$$

We shall denote this point by a bold-face letter $\mathbf{x}$ and also call it a vector $\mathbf{x}$ with components x_1, x_2, x_3 .

We have proved that addition and subtraction of vectors, as well as their multiplication by numbers, are

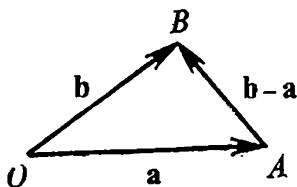


Fig. 10.

expressed in terms of triples (x_1, x_2, x_3) in the following way:

$$\left. \begin{aligned} (x_1, x_2, x_3) \pm (x'_1, x'_2, x'_3) \\ = (x_1 \pm x'_1, x_2 \pm x'_2, x_3 \pm x'_3), \\ \alpha (x_1, x_2, x_3) = (\alpha x_1, \alpha x_2, \alpha x_3). \end{aligned} \right\} \quad (12)$$

The scalar product of the vectors $\mathbf{x} = (x_1, x_2, x_3)$ and $\mathbf{x}' = (x'_1, x'_2, x'_3)$ is expressed in terms of the coordinates of the vectors $\mathbf{x}$ and $\mathbf{x}'$ by the formula

$$\mathbf{x}\mathbf{x}' = (\mathbf{x}, \mathbf{x}') = x_1x'_1 + x_2x'_2 + x_3x'_3. \quad (13)$$

The length of the vector $\mathbf{x} = (x_1, x_2, x_3)$ is a nonnegative number equal to

$$|\mathbf{x}| = \sqrt{x_1^2 + x_2^2 + x_3^2}. \quad (14)$$

Finally, the distance between the points $\mathbf{x} = (x_1, x_2, x_3)$ and $\mathbf{x}' = (x'_1, x'_2, x'_3)$ is equal to

$$|\mathbf{x} - \mathbf{x}'| = \sqrt{(x'_1 - x_1)^2 + (x'_2 - x_2)^2 + (x'_3 - x_3)^2}. \quad (15)$$

Real space whose geometry was studied here is called a *three-dimensional space*, since its points are naturally represented by triples of real numbers. We shall denote it by R_3 .

It is quite natural to denote an arbitrary plane by R_2 . In R_2 we can specify a rectangular coordinate system x_1, x_2 with the aid of which any point of R_2 or its radius vector can be represented by a pair of numbers (x_1, x_2) . It is obvious that for vectors belonging to a plane the operations of addition and subtraction, as well as of multiplication by a number, obey conditions $\{(12)\}$, provided that the third components are deleted everywhere in the parentheses. The scalar product of vectors belonging to our plane is also expressed by formula (13) where the third term of the right-hand member should be deleted. The same should be done with formulas (14) and (15).

The space R_n , where n is an arbitrary natural number, is a generalization of the spaces R_2 and R_3 . For $n > 3$ the space R_n is a mathematical fiction, or rather a great invention. It helps to mathematically understand complex phenomena.

Problems

1. Find the lengths of the vectors $(1, 1, 1)$, $(1, -1, 1)$, and $(1, 2, 3)$.
2. Find the angle between the vectors $(1, 0, 1)$ and $(1, 2, 2)$.
3. Given a unit cube (with the edge of length equal to 1). Find the angle between diagonals emanating from one of its vertices: (a) the principal diagonal and diagonal of a face; (b) two diagonals of two faces.

Sec. 6. n -Dimensional Euclidean Space.

Scalar Product of Two Vectors

The set of all possible systems $(x_1, \dots, x_n)$ of real (complex) numbers is called the n -dimensional real (complex) space and is denoted by R_n . Each system will be denoted by one (bold-face) letter without a subscript:

$$\mathbf{x} = (x_1, \dots, x_n)$$

and called a *point* or *vector* R_n . The numbers $x_1, \dots, x_n$ are termed the *coordinates* of the point (vector) $\mathbf{x}$ or the *components* of the vector $\mathbf{x}$.

Two points

$$\mathbf{x} = (x_1, \dots, x_n), \quad \mathbf{x}' = (x'_1, \dots, x'_n)$$

are regarded as equal if their corresponding coordinates are equal

$$x_j = x'_j \quad (j = 1, \dots, n).$$

In other cases $\mathbf{x}$ and $\mathbf{x}'$ are different ($\mathbf{x} \neq \mathbf{x}'$).

The systems (vectors) $\mathbf{x} = (x_1, \dots, x_n)$, $\mathbf{y} = (y_1, \dots, y_n)$ can be added and subtracted, and also multiplied by the numbers $\alpha, \beta, \dots$ which are real if R_n is a real space, and complex if R_n is a complex space*.

The *sum of vectors* $\mathbf{x}$ and $\mathbf{y}$ is defined as the vector

$$\mathbf{x} + \mathbf{y} = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n), \quad (1)$$

and their difference as the vector

$$\mathbf{x} - \mathbf{y} = (x_1 - y_1, \dots, x_n - y_n). \quad (2)$$

The *product of the number α by the vector $\mathbf{x}$* , or the vector $\mathbf{x}$ by the number α , is defined as the vector

$$\alpha \mathbf{x} = \mathbf{x} \alpha = (\alpha x_1, \dots, \alpha x_n).$$

Obviously, the following properties are fulfilled:

- (1) $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$,
- (2) $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$,
- (3) $\mathbf{x} - \mathbf{y} = \mathbf{x} + (-1) \mathbf{y}$,
- (4) $\alpha \mathbf{x} + \alpha \mathbf{y} = \alpha (\mathbf{x} + \mathbf{y})$,
- (5) $\alpha \mathbf{x} + \beta \mathbf{x} = (\alpha + \beta) \mathbf{x}$,
- (6) $\alpha (\beta \mathbf{x}) = (\alpha \beta) \mathbf{x}$,
- (7) $1 \cdot \mathbf{x} = \mathbf{x}$,

where α, β are numbers and $\mathbf{x}, \mathbf{y} \in R_n$.

The number (nonnegative)

$$|\mathbf{x}| = \sqrt{\sum_{k=1}^n |x_k|^2} = \sqrt{\sum_{k=1}^n x_k \bar{x}_k} \quad (3)$$

is called the length or norm of the vector $\mathbf{x} = (x_1, \dots, x_n)$.

* For complex numbers see our book "Differential and Integral Calculus" (Sec. 5.3), Mir Publishers, 1982 (in English).

The distance between two points $\mathbf{x} = (x_1, \dots, x_n)$ and $\mathbf{y} = (y_1, \dots, y_n)$ is determined by the formula

$$|\mathbf{x} - \mathbf{y}| = \sqrt{\sum_{k=1}^n |x_k - y_k|^2}. \quad (4)$$

The scalar product of a vector $\mathbf{x} = (x_1, \dots, x_n)$ by a vector $\mathbf{y} = (y_1, \dots, y_n)$ is the number

$$(\mathbf{x}, \mathbf{y}) = \sum_{j=1}^n x_j \bar{y}_j, \quad (5)$$

where $\bar{y}_j$ is a complex number conjugate to the number y_j .

The scalar product possesses the following properties:

(a) $(\mathbf{x}, \mathbf{x}) \geq 0$; it is equal to zero if and only if $\mathbf{x} = \mathbf{0} = (0, 0, \dots, 0)$,

(b) $\overline{(\mathbf{x}, \mathbf{y})} = (\mathbf{y}, \mathbf{x})$,

(c) $(\alpha\mathbf{x} + \beta\mathbf{y}, \mathbf{z}) = \alpha(\mathbf{x}, \mathbf{z}) + \beta(\mathbf{y}, \mathbf{z})$.

Indeed,

$$(\mathbf{x}, \mathbf{x}) = \sum_{j=1}^n x_j \bar{x}_j = \sum_{j=1}^n |x_j|^2 \geq 0,$$

where equality may be the case if $x_1 = \dots = x_n = 0$. Further,

$$\overline{(\mathbf{x}, \mathbf{y})} = \overline{\left(\sum_{j=1}^n x_j \bar{y}_j\right)} = \sum_{j=1}^n \bar{x}_j y_j = \sum_{j=1}^n y_j \bar{x}_j = (\mathbf{y}, \mathbf{x}).$$

Here we have taken advantage of the operation of conjugation $\overline{z + z_1} = \bar{z} + \bar{z}_1$ and $\overline{zz_1} = \bar{z} \cdot \bar{z}_1$. Finally,

$$\begin{aligned} (\alpha\mathbf{x} + \beta\mathbf{y}, \mathbf{z}) &= \sum_{j=1}^n (\alpha x_j + \beta y_j) \bar{z}_j \\ &= \alpha \sum_{j=1}^n x_j \bar{z}_j + \beta \sum_{j=1}^n y_j \bar{z}_j = \alpha(\mathbf{x}, \mathbf{z}) + \beta(\mathbf{y}, \mathbf{z}). \end{aligned}$$

We shall confine ourselves to the consideration of a real n -dimensional space. In it vectors are multiplied

* The space R_n into which a scalar product is introduced is called *Euclidean*.

by real numbers. The term "real" will sometimes be omitted. For a real space the modulus sign (of numbers) in formulas (3) and (4) may be substituted by parentheses:

$$|\mathbf{x}| = \sqrt{\sum_{j=1}^n x_j^2}, \quad (3')$$

$$|\mathbf{x} - \mathbf{y}| = \sqrt{\sum_{j=1}^n (x_j - y_j)^2}, \quad (4')$$

and the scalar product has the form

$$(\mathbf{x}, \mathbf{y}) = \sum_{j=1}^n x_j y_j, \quad (5')$$

since in this case $\bar{y}_j = y_j$, and then $(\mathbf{x}, \mathbf{y}) = (\mathbf{y}, \mathbf{x})$, i.e. *in a real space the scalar product is commutative.*

From the properties (a), (b), (c) there follows an important inequality (called *Buniakovski's inequality**)

$$|(\mathbf{x}, \mathbf{y})| \leq \sqrt{(\mathbf{x}, \mathbf{x})} \sqrt{(\mathbf{y}, \mathbf{y})}. \quad (6)$$

Indeed, for any real number λ

$$\begin{aligned} 0 &\leq (\mathbf{x} + \lambda \mathbf{y}, \mathbf{x} + \lambda \mathbf{y}) \\ &= (\mathbf{x}, \mathbf{x}) + \lambda (\mathbf{y}, \mathbf{x}) + \lambda (\mathbf{x}, \mathbf{y}) + \lambda^2 (\mathbf{y}, \mathbf{y}) \\ &= (\mathbf{x}, \mathbf{x}) + 2 (\mathbf{x}, \mathbf{y}) \lambda + (\mathbf{y}, \mathbf{y}) \lambda^2 = a + 2b\lambda + c\lambda^2, \end{aligned}$$

where $a = (\mathbf{x}, \mathbf{x})$, $b = (\mathbf{x}, \mathbf{y})$, $c = (\mathbf{y}, \mathbf{y})$. We see that the quadratic polynomial

$$a + 2b\lambda + c\lambda^2 \geq 0 \quad (-\infty < \lambda < \infty)$$

is nonnegative for any real λ . Therefore its entire graph lies above the axis λ which is possible only if the discriminant of the polynomial is negative or zero, i.e. if $b^2 - ac \leq 0$ or $b^2 \leq ac$, and we have obtained Buniakovski's inequality (6).

In terms of the components of the vectors $\mathbf{x}$ and $\mathbf{y}$ inequality (6) can be written as follows:

$$\left| \sum_{j=1}^n x_j y_j \right| \leq \sqrt{\sum_{j=1}^n x_j^2} \sqrt{\sum_{j=1}^n y_j^2}. \quad (7)$$

* Buniakovski (or Bouniakowsky), Victor Jakowlevitsch (1804-1889). Russian mathematician.

Thus, whatever the real numbers x_j, y_j are, inequality (7) is fulfilled.

By virtue of (3), Buniakovski's inequality can be written in the following manner:

$$|(\mathbf{x}, \mathbf{y})| \leq |\mathbf{x}| |\mathbf{y}|.$$

But then there exists a unique number λ satisfying the inequalities $-1 \leq \lambda \leq 1$, for which the rigorous equality

$$(\mathbf{x}, \mathbf{y}) = \lambda |\mathbf{x}| |\mathbf{y}|$$

holds true.

Note that on $[0, \pi]$ the function $\cos t$ has a single-valued strictly decreasing inverse function with the domain of values on $[-1, 1]$. Therefore, for every λ ($-1 \leq \lambda \leq 1$) there exists a unique angle ω ($0 \leq \omega \leq \pi$) such that $\lambda = \cos \omega$. Hence, we have proved the equation

$$(\mathbf{x}, \mathbf{y}) = |\mathbf{x}| |\mathbf{y}| \cos \omega. \quad (8)$$

The number ω is called the *angle between n -dimensional vectors $\mathbf{x}$ and $\mathbf{y}$* , although, in fact, for $n > 3$ the vectors $\mathbf{x}$ and $\mathbf{y}$ are not real line-segments at all, they are mathematical abstractions.

The vectors $\mathbf{x}$ and $\mathbf{y}$ are called *orthogonal* if their scalar product is equal to zero:

$$(\mathbf{x}, \mathbf{y}) = \sum_{j=1}^n x_j y_j = 0.$$

It follows from (8) that for nonzero vectors $\mathbf{x}$ and $\mathbf{y}$ to be orthogonal, it is necessary and sufficient that the angle ω between them be $\pi/2$.

We should like to mention here an important inequality (called *Minkowski's inequality**)

$$|\mathbf{x} + \mathbf{y}| \leq |\mathbf{x}| + |\mathbf{y}|, \quad (9)$$

* Minkowski, Hermann (1864-1909). Number theorist, algebraist, analyst, and geometer. Born in Russia, lived in Switzerland and Germany. Created the geometry of numbers. Provided the four-dimensional geometric space-time framework for relativity theory.

or in terms of components

$$\sqrt{\sum_{j=1}^n (x_j + y_j)^2} \leq \sqrt{\sum_{j=1}^n x_j^2} + \sqrt{\sum_{j=1}^n y_j^2}. \quad (10)$$

It can be proved in the following way:

$$|\mathbf{x} + \mathbf{y}|^2 = (\mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y}) = (\mathbf{x}, \mathbf{x}) + 2(\mathbf{x}, \mathbf{y}) + (\mathbf{y}, \mathbf{y}).$$

Using Buniakovski's inequality (6), we have

$$\begin{aligned} |\mathbf{x} + \mathbf{y}|^2 &\leq (\mathbf{x}, \mathbf{x}) + 2\sqrt{(\mathbf{x}, \mathbf{x})}\sqrt{(\mathbf{y}, \mathbf{y})} + (\mathbf{y}, \mathbf{y}) \\ &= (\sqrt{(\mathbf{x}, \mathbf{x})} + \sqrt{(\mathbf{y}, \mathbf{y})})^2, \end{aligned}$$

whence follows (9). From (9) there follows the inequality

$$||\mathbf{x}| - |\mathbf{y}|| \leq |\mathbf{x} - \mathbf{y}|, \quad (11)$$

since

$$\begin{aligned} |\mathbf{x}| &= |\mathbf{x} - \mathbf{y} + \mathbf{y}| \leq |\mathbf{x} - \mathbf{y}| + |\mathbf{y}|, \\ |\mathbf{y}| &= |\mathbf{y} - \mathbf{x} + \mathbf{x}| \leq |\mathbf{x} - \mathbf{y}| + |\mathbf{x}|. \end{aligned}$$

Problem. Find the angle between the vectors (1, 0, 1, 0) and (1, 1, 1, 1).

Sec. 7. Segment of a Line. Dividing a Segment in a Given Ratio

Let us specify arbitrary points $\mathbf{x}, \mathbf{y} \in R_n$ and introduce the set of points (vectors)

$$\mathbf{z} = \lambda\mathbf{x} + \mu\mathbf{y} \quad (\lambda, \mu \geq 0, \quad \lambda + \mu = 1), \quad (1)$$

defined by nonnegative numbers λ, μ whose sum is equal to 1. We have

$$\mathbf{z} = (1 - \mu)\mathbf{x} + \mu\mathbf{y} = \mathbf{x} + \mu(\mathbf{y} - \mathbf{x}) \quad (0 \leq \mu \leq 1) \quad (2)$$

or

$$\mathbf{z} = \mathbf{y} + \lambda(\mathbf{x} - \mathbf{y}) \quad (0 \leq \lambda \leq 1). \quad (2')$$

It is seen from equation (2) that in a three-dimensional space points $\mathbf{z}$ fill the line segment connecting $\mathbf{x}$ and $\mathbf{y}$. The radius vector $\mathbf{z}$ is the sum of the vector $\mathbf{x}$ and the vector $\mu(\mathbf{y} - \mathbf{x})$ which is collinear with $\mathbf{y} - \mathbf{x}$ (Fig. 11).

Hence, the set of points (1) represents a line segment $[x, y]$ in R_3 joining the points x and y . For $\mu = 0$ $z = x$, for $\lambda = 0$ $z = y$, for any $\lambda > 0$ ($\mu = 1 - \lambda > 0$) z is an arbitrary point on $[x, y]$.

The line segment $[x, y]$, joining the points $x, y \in R_n$, is defined as the set of all points z of form (1). The following theorem holds true.

Theorem 1. *The point*

$$z = \lambda x + \mu y \quad (\lambda, \mu \geq 0, \quad \lambda + \mu = 1)$$

divides the segment $[x, y]$ joining the points $x, y \in R_n$ into

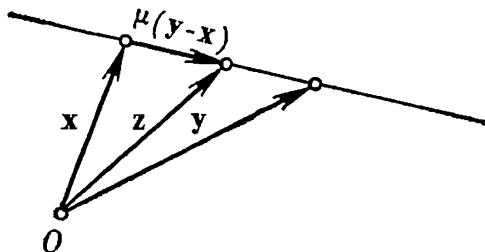


Fig. 11.

two segments whose lengths are in the ratio $\mu : \lambda$.

Proof. It follows from (2) that $z - x = \mu (y - x)$, and therefore the distance between the points x and z is equal to

$$|z - x| = \mu |y - x|. \quad (3)$$

Further from (2'): $z - y = \lambda (x - y)$, and therefore the distance between the points z and y is equal to

$$|z - y| = \lambda |x - y|. \quad (4)$$

It follows from (3) and (4):

$$|z - x| : |z - y| = \mu : \lambda,$$

which was to be proved.

Problem. On the segment $[x, y]$, joining the points $x, y \in R_n$, find the point z that divides this segment in the ratio $p : q$ ($p > 0, q > 0$).

Solution. Let us take the numbers

$$\lambda = \frac{q}{p+q}, \quad \mu = \frac{p}{p+q}.$$

They satisfy the properties of $\lambda, \mu \geq 0, \lambda + \mu = 1, \mu/\lambda = p/q$. Therefore, by Theorem 1, the required point

$$\mathbf{z} = \lambda \mathbf{x} + \mu \mathbf{y} = \frac{q\mathbf{x} + p\mathbf{y}}{p+q}. \quad (5)$$

Its coordinates $\mathbf{z} = (z_1, \dots, z_n)$ are expressed in terms of the coordinates $\mathbf{x} = (x_1, \dots, x_n), \mathbf{y} = (y_1, \dots, y_n)$ by the equation

$$z_j = \frac{qx_j + py_j}{p+q} \quad (j = 1, \dots, n). \quad (5')$$

In particular, the middle of the segment is obtained for $p = q = 1$, i.e. $\lambda = \mu = 1/2$.

Note, that, as it is proved in mechanics, the point $\mathbf{z}$ determined by (5) or (5') is the centre of gravity of the system of points $\mathbf{x}$ and $\mathbf{y}$ at which masses q and p are concentrated, respectively.

Note that in R_3 the set of points

$$\mathbf{z} = \lambda \mathbf{x} + \mu \mathbf{y}, \quad \lambda + \mu = 1,$$

where λ and μ are of either sign, represents a straight line passing through the points $\mathbf{x}$ and $\mathbf{y}$. It is obvious from (2').

And in space R_n ($n > 3$) this set is called a straight line by definition.

Example 1. Find coordinates of the centre of gravity of a system of material points $\mathbf{x}^k = (x_1^k, x_2^k, x_3^k)$ with corresponding masses p_k ($k = 1, \dots, M$). Using formula (5') for the points $\mathbf{x}^1, \mathbf{x}^2$ we first find the centre of gravity $\mathbf{z}^1$ of the points $\mathbf{x}^1$ and $\mathbf{x}^2$, and then the centre of gravity $\mathbf{z}^2$ of the points $\mathbf{z}^1$ and $\mathbf{x}^3$ with corresponding masses $p_1 + p_2$ and p_3 . Proceeding in the same manner, we obtain for the $(M - 1)$ st step

$$z_j^{M-1} = \frac{1}{p_1 + \dots + p_M} \sum_{k=1}^M p_k x_j^k \quad (j = 1, 2, 3).$$

Sec. 8. The Straight Line

The concept of the straight line is primary in geometry. From the axioms of geometry we know that a straight line is determined by two points, i.e. through two given points we can draw only one straight line. Through a point lying

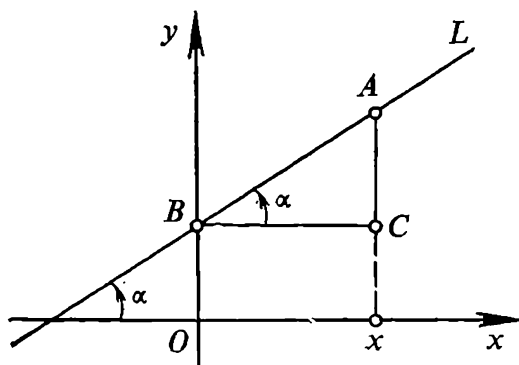


Fig. 12.

on a given straight line we can draw only one straight line perpendicular to the given line.

In the plane, let us specify a rectangular coordinate system (x, y) and a straight line L not parallel to the y -axis (Fig. 12). We know from school that the equation of the straight line L has the form

$$y = kx + l, \quad (1)$$

where $k = \tan \alpha$ and α is the angle formed by the line L with the positive direction of the x -axis, and l is the ordinate of the point of intersection of L with the y -axis ($l = OB$).

When we say that equation (1) is the equation of the straight line L , we mean by this that L is a locus of points whose coordinates (x, y) satisfy equation (1). The validity of this assertion is obvious from Fig. 12. Point A is an arbitrary (moving) point belonging to the line L ; its coordinates (x, y) are: $BC = x$, $AC = y - l$, and

$$k = \tan \alpha = \frac{y-l}{x}, \quad (1')$$

whence (1) follows. Conversely, equation (1) is equivalent to equation (1'), and the latter, obviously, expresses the fact that the point (x, y) lies on the straight line L . In Fig. 12 the angle α is acute. In the case of an obtuse α we reason in a similar way.

Consider the equation

$$Ax + By + C = 0, \quad (2)$$

where A, B, C are specified numbers, A and B being not both zero.

If $B \neq 0$, then equation (2) can be written in the form

$$y = -\frac{A}{B}x - \frac{C}{B}, \quad (2')$$

or, setting

$$k = -\frac{A}{B}, \quad l = -\frac{C}{B},$$

in form (1). Since equations (2) and (2') are equivalent to each other, i.e. any point (x, y) , satisfying one of them, satisfies the other, (2) for $B \neq 0$ is the equation of a straight line inclined to the positive direction of the x -axis at an angle α whose tangent equals $-A/B$ ($\tan \alpha = -A/B$), and intersecting the y -axis at the point having the ordinate $-C/B$ ($l = -C/B$). For $B = 0$ equation (2) takes the form

$$Ax + C = 0 \quad (A \neq 0!),$$

or

$$x = a \quad (a = -C/A).$$

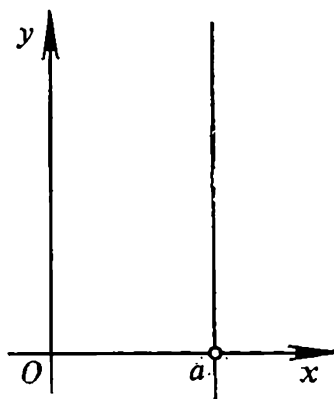


Fig. 13.

This is also an equation of a straight line which is parallel to the y -axis. Namely, this is a locus of points (x, y) the abscissas x of which are equal to one and the same number a . Figure 13 shows such a line for $a > 0$.

It follows from the aforesaid that (2), where A, B, C are given numbers, A and B are not both zero, is the

equation of some straight line. For $B \neq 0$ this line is not parallel to the y -axis. In particular, for $A = 0$ it is parallel to the x -axis ($y = -C/B$). If $B = 0$, then it is parallel to the y -axis. Note that $y = 0$ is the equation of the x -axis, and $x = 0$ is the equation of the y -axis.

Equation (2) is called the *general equation of the straight line*. Any straight line situated arbitrarily with respect to the coordinate system can be described by the equation of form (2) for suitable constant numbers A , B , C . We underline once again that the *numbers A and B in equation (2) are not both zero*. Note that the number k in equation (1) is called the *slope of a straight line*.

Let us solve several important problems.

Problem 1. Write the equation of a straight line, whose slope is equal to the number k , passing through a given point (x_0, y_0) .

Solution. A straight line with the slope k has the form

$$y = kx + l, \quad (3)$$

where l may be any number. Since the point (x_0, y_0) must lie on the given line, its coordinates must satisfy the equation

$$y_0 = kx_0 + l. \quad (4)$$

Subtracting (4) from (3) we get the required equation

$$y - y_0 = k(x - x_0) \quad (5)$$

of the straight line passing through the point (x_0, y_0) and having the slope k .

Problem 2. Write the equation of a straight line passing through two given points (x_1, y_1) and (x_2, y_2) . It is assumed that the points are different.

Solution. Let $x_1 \neq x_2$. Then, obviously, the sought-for line is not parallel to the y -axis, and therefore may be written in the form

$$y - y_1 = k(x - x_1), \quad (6)$$

where k is a certain number. Equation (6) itself expresses the fact that the straight line passes through the point (x_1, y_1) . For this line to pass also through the point (x_2, y_2) , it is necessary that the equation

$$y_2 - y_1 = k(x_2 - x_1) \quad (7)$$

be fulfilled. Dividing (6) by (7) (that is, dividing the left-hand member of (6) by that of (7), and the right-hand member of (6) by that of (7)), we get

$$\frac{y - y_1}{y_2 - y_1} = \frac{x - x_1}{x_2 - x_1}. \quad (8)$$

This is just the equation of the straight line passing through the points (x_1, y_1) and (x_2, y_2) .

Remark. It might happen that $y_2 - y_1 = 0$, then, formally, we would obtain the equation

$$\frac{y - y_1}{0} = \frac{x - x_1}{x_2 - x_1}.$$

In spite of the fact that this equation is senseless, for the sake of convenience, it is customarily written in such a way. Getting rid of the denominators, we obtain a true equality

$$y - y_1 = 0 \cdot \frac{x - x_1}{x_2 - x_1} = 0$$

or

$$y = y_1. \quad (9)$$

The case $x_1 = x_2 = c$ leads to the solution $x = c$.

Problem 3. Find the angle ω between the straight lines

$$y = k_1x + l_1, \quad y = k_2x + l_2.$$

Solution. We have $k_1 = \tan \alpha_1$, $k_2 = \tan \alpha_2$, where α_1 , α_2 are the angles formed by the given lines with the positive direction of the x -axis, respectively. We have (Fig. 14)

$$\tan \omega = \tan (\alpha_1 - \alpha_2) = \frac{\tan \alpha_1 - \tan \alpha_2}{1 + \tan \alpha_1 \tan \alpha_2} = \frac{k_1 - k_2}{1 + k_1 k_2}, \quad (10)$$

which is the formula for the angle between two straight lines.

The case $1 + k_1 k_2 = 0$ or

$$k_1 \cdot k_2 = -1 \quad (11)$$

expresses the condition for perpendicularity of straight lines. The condition for parallelism of straight lines ($\tan \omega = 0$) will be written as follows:

$$k_1 = k_2. \quad (12)$$

Let us write equation (2) in the following form:

$$A_1x_1 + A_2x_2 + C = 0, \quad (2')$$

where we have set $x = x_1$, $y = x_2$, $A = A_1$, $B = A_2$. For $A_1 \neq 0$, $A_2 \neq 0$, $C \neq 0$ equation (2') can be written in the form

$$\frac{x_1}{a} + \frac{x_2}{b} = 1, \quad a = \frac{-C}{A_1}, \quad b = \frac{-C}{A_2}. \quad (13)$$

Equation (13) is the *intercept form of the equation of a straight line*. This line intersects the axis x_1 (the

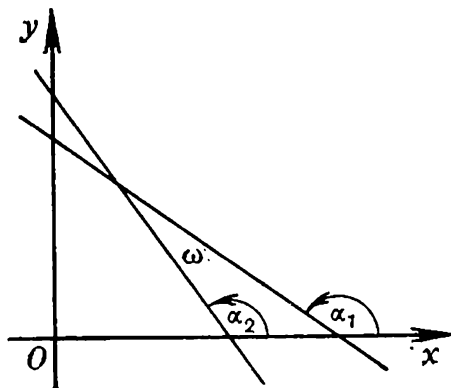


Fig. 14.

straight line $x_2 = 0$) at the point $(a, 0)$ and the axis x_2 at the point $(0, b)$.

If a straight line, satisfying equation (2'), passes through the point (x_1^0, x_2^0) , then

$$A_1x_1^0 + A_2x_2^0 + C = 0. \quad (14)$$

Subtracting (14) from (2'), we get

$$A_1(x_1 - x_1^0) + A_2(x_2 - x_2^0) = 0. \quad (15)$$

Equation (15) is the *equation of a straight line passing through a point (x_1^0, x_2^0)* .

Introducing into consideration the vectors $\mathbf{A} = (A_1, A_2)$, $\mathbf{x} = (x_1, x_2)$, $\mathbf{x}^0 = (x_1^0, x_2^0)$, we may regard the left-hand member of (15) as a scalar product of the vector $\mathbf{A}$ by the vector $\mathbf{x} - \mathbf{x}^0$. Therefore equation (15) can be written

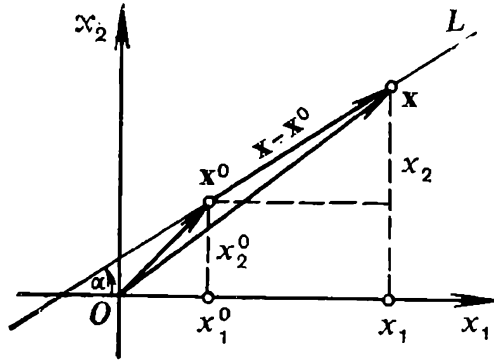


Fig. 15.

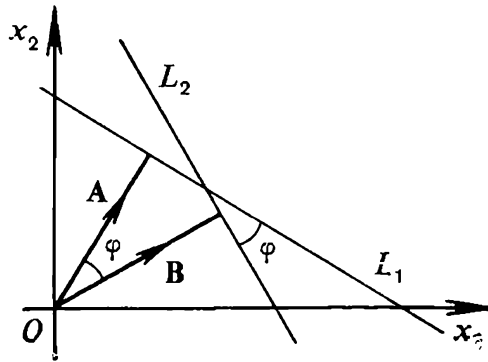


Fig. 16.

in the vector form as

$$\mathbf{A}(\mathbf{x} - \mathbf{x}^0) = 0. \quad (15')$$

The vector $\mathbf{x} - \mathbf{x}^0$ belongs to the line L (Fig. 15). Hence, it is seen from (15') that the vector $\mathbf{A} = (A_1, A_2)$ is orthogonal (perpendicular) to the given line. Thus, we have elucidated the geometrical meaning of the coefficients A_1 and A_2 .

Consider two lines

$$A_1x_1 + A_2x_2 + C_1 = 0 \quad (L_1), \quad (16)$$

$$B_1x_1 + B_2x_2 + C_2 = 0 \quad (L_2). \quad (17)$$

Since the vectors $\mathbf{A} = (A_1, A_2)$ and $\mathbf{B} = (B_1, B_2)$ are perpendicular to lines (16) and (17), respectively, the angle φ between the straight lines (16) and (17) is equal to the angle between the vectors $\mathbf{A}$ and $\mathbf{B}$ (Fig. 16). The angle φ

can be computed by the formula

$$\cos \varphi = \frac{(\mathbf{A}, \mathbf{B})}{|\mathbf{A}| |\mathbf{B}|} = \frac{A_1 B_1 + A_2 B_2}{\sqrt{A_1^2 + A_2^2} \sqrt{B_1^2 + B_2^2}}. \quad (18)$$

Remark. If φ is an angle between two straight lines, then $\pi - \varphi$ is also an angle between these lines. The number (18) can be either positive, or negative. The former corresponds to the angle φ , the latter to $\pi - \varphi$.

From (18) we obtain the *condition for the perpendicularity* of L_1 and L_2 ($\varphi = \pi/2$, $\cos \varphi = 0$):

$$A_1 B_1 + A_2 B_2 = 0. \quad (19)$$

If the lines L_1 and L_2 are *parallel*, then the vectors $\mathbf{A}$ and $\mathbf{B}$ are collinear and $\mathbf{A} = \lambda \mathbf{B}$, where λ is some real

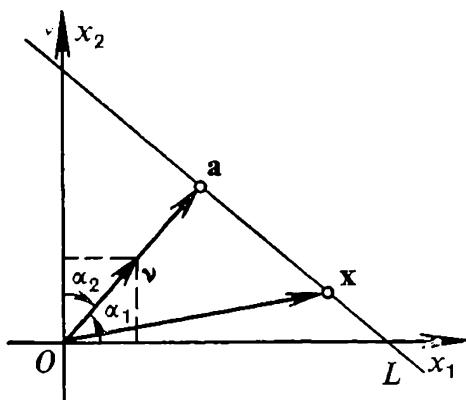


Fig. 17.

number. Hence the condition for parallelism of straight lines is expressed by the equation

$$\frac{A_1}{B_1} = \frac{A_2}{B_2}. \quad (20)$$

Let there be given an arbitrary straight line L in a rectangular coordinate system (Fig. 17) not passing through the origin, and let $\mathbf{a}$ be a vector emanating from the origin and perpendicular to the line L with the foot lying on this line. The vector $\mathbf{a}$ completely determines the line L , since through its foot there passes the only straight line perpendicular to it. Let p be the length of the vector $\mathbf{a} = (p = |\mathbf{a}|)$; $\mathbf{v} = (\cos \alpha_1, \cos \alpha_2)$ is a unit vector in the

same direction as $\mathbf{a}$. Here α_i is the angle between $\mathbf{a}$ (or $\mathbf{v}$) and the positive direction of the axis x_i ($i = 1, 2$); $\cos^2 \alpha_1 + \cos^2 \alpha_2 = 1$ ($\cos \alpha_2 = \sin \alpha_1$). Let us denote by $\mathbf{x} = (x_1, x_2)$ an arbitrary (moving) point of the line L , or, which is the same, its radius vector. The projection of the vector $\mathbf{x}$ on the unit vector $\mathbf{v}$ is, obviously, equal to p , i.e. the scalar product of the radius vector of an arbitrary point $\mathbf{x}$ on the line L by the vector $\mathbf{v}$ is equal to p :

$$(\mathbf{x}, \mathbf{v}) = |\mathbf{v}| \operatorname{pr}_{\mathbf{v}} \mathbf{x} = p. \quad (21)$$

Thus, we have obtained a vector equation of L , since, conversely, if a point (radius vector) $\mathbf{x}$ satisfies equation (21), then it lies on L (a point not lying on L has a projection of $\mathbf{v}$ which is different from p).

If L passes through the origin, then its equation can also be written in the form of (21), where $\mathbf{v}$ is a unit vector perpendicular to it.

In the coordinate form, equation (21) is written as follows:

$$x_1 \cos \alpha_1 + x_2 \cos \alpha_2 = p \quad (p \geq 0) \quad (21')$$

or

$$x_1 \cos \alpha_1 + x_2 \sin \alpha_1 = p \quad (p \geq 0). \quad (21'')$$

This is the *normal form of the equation of a straight line*.

If a straight line L is given by the general equation (2')

$$A_1 x_1 + A_2 x_2 + C = 0,$$

then it can be reduced to the normal form by multiplying it by the number

$$M = \pm 1 / \sqrt{A_1^2 + A_2^2}, \quad (22)$$

where the sign should be opposite to the sign of C ($p = -MC \geq 0$). The number M is called the *normalizing factor*. Since

$$(MA_1)^2 + (MA_2)^2 = 1,$$

there exists the only angle α_1 satisfying the inequality $0 \leq \alpha_1 < 2\pi$, for which

$$MA_1 = \cos \alpha_1, \quad MA_2 = \sin \alpha_1. \quad (23)$$

As a result, we obtain equation (21), where $p = -MC \geq 0$. Note once again that the number p is equal to the distance from the origin to the straight line.

Problem. Find the distance d from the point $\mathbf{x}^0 = (x_1^0, x_2^0)$ to the line L defined by the equation

$$A_1x_1 + A_2x_2 + C = 0. \quad (24)$$

Solution. Let

$$(\mathbf{x}, \mathbf{v}) - p = 0 \quad (25)$$

be a normal equation of line (24). Thus, if $C \neq 0$, then p ($p > 0$) is the length of the vector $\vec{OA}$ dropped from the origin O on L and $\mathbf{v}$ is a unit vector which is in the same direction as $\vec{OA}$ ($p = |\vec{OA}|$, $\mathbf{v} = \vec{OA}/p$, see

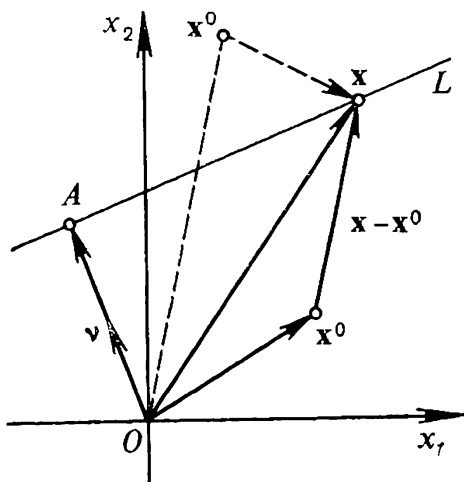


Fig. 18

Fig. 18). Let $\mathbf{x}$ be an arbitrary point on L . Then, obviously, in order to find the distance from the point $\mathbf{x}^0$ to L , we have to project the vector $\mathbf{x} - \mathbf{x}^0$ on the direction of the vector $\mathbf{v}$ and take the absolute value of the projection thus obtained:

$$\begin{aligned} d &= |\text{pr}_{\mathbf{v}} (\mathbf{x} - \mathbf{x}^0)| = |(\mathbf{x} - \mathbf{x}^0, \mathbf{v})| \\ &= |(\mathbf{x}, \mathbf{v}) - (\mathbf{x}^0, \mathbf{v})| = |p - (\mathbf{x}^0, \mathbf{v})| \\ &= |(\mathbf{x}^0, \mathbf{v}) - p|. \end{aligned}$$

We have obtained the formula

$$d = |(\mathbf{x}^0, \mathbf{v}) - p|. \quad (26)$$

Thus, in order to get the distance d , we have to reduce equation (24) to the normal form, transpose p to its left-

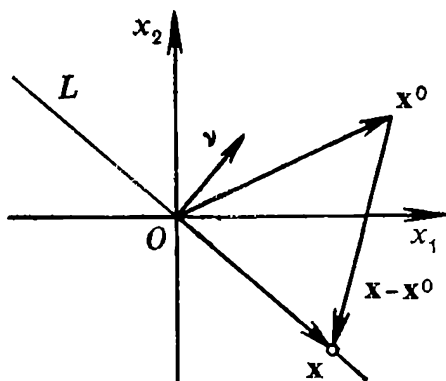


Fig. 19

hand side, replace x_1, x_2 by the corresponding coordinates x_1^0, x_2^0 of the point $\mathbf{x}^0$, and take the absolute value of the expression thus obtained.

This is how formula (26) is written in terms of the coefficients A_1, A_2, C :

$$d = \frac{|A_1 x_1^0 + A_2 x_2^0 + C|}{\sqrt{A_1^2 + A_2^2}}. \quad (26')$$

For $C = 0$ formulas (26)-(26') also hold true. In this case $p = 0$, $\mathbf{v}$ is one of the two unit vectors perpendicular to L (Fig. 19). Now

$$\begin{aligned} d &= |\text{pr}_{\mathbf{v}}(\mathbf{x} - \mathbf{x}^0)| = |(\mathbf{v}, \mathbf{x} - \mathbf{x}^0)| \\ &= |(\mathbf{v}, \mathbf{x}) - (\mathbf{v}, \mathbf{x}^0)| = |(\mathbf{v}, \mathbf{x}^0)| \end{aligned}$$

or

$$d = \frac{|A_1 x_1^0 + A_2 x_2^0|}{\sqrt{A_1^2 + A_2^2}},$$

i.e. (26') for $C = 0$.

Remark. It is obvious from Fig. 18 that if the origin and the point $\mathbf{x}^0$ are found (a) on one side of L , then the

angle between $\mathbf{v}$ and $\mathbf{x} - \mathbf{x}^0$ is acute, and $d = p - (\mathbf{x}^0, \mathbf{v})$, (b) on different sides of L , then the angle between $\mathbf{x} - \mathbf{x}^0$ and $\mathbf{v}$ is an obtuse angle and $d = (\mathbf{x}^0, \mathbf{v}) - p$.

Sec. 9. The Equation of a Plane

In space R_3 , with a rectangular coordinate system (x_1, x_2, x_3) introduced, let us specify a vector $\mathbf{a}$ emanating from the origin O . Through its terminus a we pass a plane L perpendicular to $\mathbf{a}$ (Fig. 20) and denote an arbitrary (moving) point of the plane L by $\mathbf{x} = (x_1, x_2, x_3)$. The letter $\mathbf{x}$ also denotes the radius vector of the point $\mathbf{x}$.

Let $p = |\mathbf{a}|$ be the length of the vector $\mathbf{a}$, and

$$\mathbf{v} = (\cos \alpha_1, \cos \alpha_2, \cos \alpha_3)$$

be a unit vector which is in the same direction as $\mathbf{a}$. Here, $\alpha_1, \alpha_2, \alpha_3$ are angles made by the vector $\mathbf{v}$ with the positive

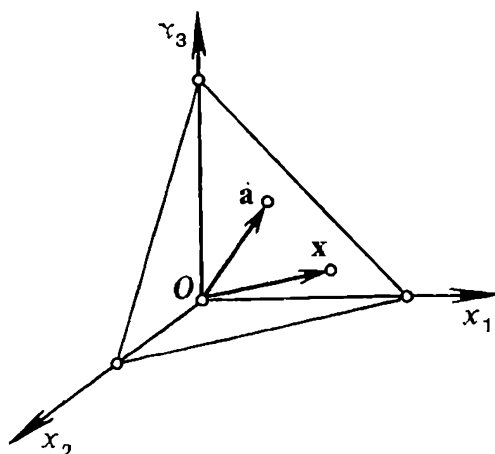


Fig. 20

directions of the axes x_1, x_2, x_3 , respectively. The projection of any point $\mathbf{x} \in L$ on the vector $\mathbf{v}$ is, obviously, a constant quantity equal to p :

$$(\mathbf{x}, \mathbf{v}) = p \quad (p \geq 0). \quad (1)$$

Equation (1) has sense for $p = 0$ as well. In this case plane L passes through the origin O ($a = 0$), $\mathbf{v}$ being a unit vector drawn from O perpendicular to L in either direction, i.e. the vector $\mathbf{v}$ is defined up to a sign. Equation (1) is the *equation of plane L in a vector form*. In terms of coordinates it is written as

$$x_1 \cos \alpha_1 + x_2 \cos \alpha_2 + x_3 \cos \alpha_3 = p \quad (p \geq 0) \quad (1')$$

and is called the *normal form of the equation of a plane*.

Multiplying equation (1') by a nonzero number, we get an equivalent equation of the form

$$A_1x_1 + A_2x_2 + A_3x_3 + B = 0, \quad (2)$$

defining the same plane. Here the numbers A_1, A_2, A_3 are not all zero. Equation (2), where the numbers A_1, A_2, A_3 are not all zero, is called the *general equation of a plane*.

An arbitrary equation of form (2), where the numbers A_1, A_2, A_3 are not all zero, can be reduced to the normal form by multiplying it by the number

$$M = \pm 1 / \sqrt{A_1^2 + A_2^2 + A_3^2},$$

where the sign is taken opposite to that of the number B . Then the number $p = -MB$ will be nonnegative and equation (2) is transformed into the following equivalent equation:

$$MA_1x_1 + MA_2x_2 + MA_3x_3 = p. \quad (3)$$

Here

$$(MA_1)^2 + (MA_2)^2 + (MA_3)^2 = 1.$$

This shows that the vector

$$\mathbf{v} = (MA_1, MA_2, MA_3)$$

is a unit one ($|\mathbf{v}| = 1$). Its projections on the coordinate axes are equal to

$$MA_1 = \cos \alpha_1,$$

$$MA_2 = \cos \alpha_2,$$

$$MA_3 = \cos \alpha_3.$$

where $\alpha_1, \alpha_2, \alpha_3$ are angles formed by the vector $\mathbf{v}$ with the positive directions of the axes x_1, x_2, x_3 , respectively. By virtue of the introduced notation, equation (3) has the form

$$x_1 \cos \alpha_1 + x_2 \cos \alpha_2 + x_3 \cos \alpha_3 = p \quad (p \geq 0), \quad (3')$$

i.e. we have obtained the normal form of the equation of a plane.

If we are given general equation of a plane (2) and it is required to find out its position relative to the coordinate system, then it is sufficient to reduce equation (2) to the normal form by multiplying it by the normalizing factor M .

From equation (2) itself (without carrying out any computations) we can only state the following two facts: (1) if $B = 0$, then the plane passes through the origin, and if $B \neq 0$, then it does not pass through the origin; (2) the vector $\mathbf{A} = (A_1, A_2, A_3)$ is perpendicular to the plane, since it is collinear with the unit vector $\mathbf{v} = (MA_1, MA_2, MA_3)$, perpendicular to the given plane.

The equation

$$A_1x_1 + A_2x_2 + B = 0 \quad (4)$$

is a particular case of equation (2). In the plane (x_1, x_2) equation (4) defines a straight line, and in space (x_1, x_2, x_3) it is the equation of the plane L perpendicular to the coordinate plane (x_1, x_2) and passing through this line. Whatever the point (x_1, x_2, x_3) , belonging to the plane L , may be, its coordinates x_1, x_2 satisfy equation (4) irrespective of its third coordinate x_3 . The equation

$$A_1x_1 + B = 0 \quad (A_1 \neq 0) \quad (5)$$

is a particular case of equation (4). It can also be written in the form

$$x_1 = C \quad (C = -B/A_1). \quad (5')$$

In space (x_1, x_2, x_3) , equation (5') is a locus of points (x_1, x_2, x_3) whose first coordinate is equal to C , the coordinates x_2, x_3 being equal to any numbers. It is clear that (5') determines a plane parallel to the coordinate plane x_2, x_3 (or perpendicular to the axis x_1).

Intercept form of the equation of a plane. If $A_i = 0$ ($i = 1, 2, 3$) and $B \neq 0$, then equation (2) can be written in the following way:

$$\frac{x_1}{a_1} + \frac{x_2}{a_2} + \frac{x_3}{a_3} = 1, \quad (6)$$

where $a_i = -B/A_i$ ($i = 1, 2, 3$). Equation (6) is called the *intercept form of the equation of a plane*. This plane (see Fig. 21) intersects the axis Ox_1 at the point $(a_1, 0, 0)$,

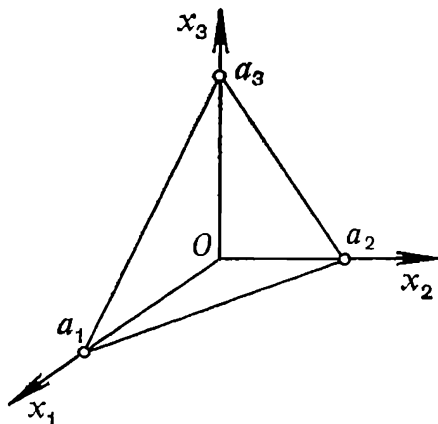


Fig. 21

the axis Ox_2 at the point $(0, a_2, 0)$, and the axis Ox_3 at the point $(0, 0, a_3)$. Judging by equation (6), it is easy to imagine the position of the plane relative to the coordinate system.

The equation of a plane passing through a given point. If the point $\mathbf{x}^0 = (x_1^0, x_2^0, x_3^0)$ lies in plane (2), then its coordinates satisfy equation (2):

$$\sum_{i=1}^3 A_i x_i^0 + B = 0. \quad (7)$$

Subtracting (7) from (2), we get

$$\sum_{i=1}^3 A_i (x_i - x_i^0) = 0. \quad (8)$$

Equation (8) is known as the *equation of a plane passing through the point* $\mathbf{x}^0 = (x_1^0, x_2^0, x_3^0)$. In terms of vectors,

equation (8) has the form

$$\mathbf{A} (\mathbf{x} - \mathbf{x}^0) = 0. \quad (8')$$

Since the vector $\mathbf{x} - \mathbf{x}^0$, applied at the point $\mathbf{x}^0$, belongs to plane (1), equation (8') implies that the vector $\mathbf{A}$ is orthogonal to plane (1), which was ascertained before from other reasonings.

The equation of a plane passing through three given points. Let there be given three noncollinear points (i.e. not lying on a straight line)

$$\begin{aligned} \mathbf{x}^1 &= (x_1^1, x_2^1, x_3^1), \\ \mathbf{x}^2 &= (x_1^2, x_2^2, x_3^2), \\ \mathbf{x}^3 &= (x_1^3, x_2^3, x_3^3). \end{aligned}$$

It is required to write the equation of the plane passing through these three points. We know from geometry that there exists such a plane which is unique. Since it passes through the point $\mathbf{x}^1$, its equation has the form

$$A_1 (x_1 - x_1^1) + A_2 (x_2 - x_2^1) + A_3 (x_3 - x_3^1) = 0, \quad (9)$$

where A_1, A_2, A_3 are not all zero. Since it also passes through the points $\mathbf{x}^2, \mathbf{x}^3$, the following conditions must be fulfilled:

$$\left. \begin{aligned} A_1 (x_1^2 - x_1^1) + A_2 (x_2^2 - x_2^1) + A_3 (x_3^2 - x_3^1) &= 0, \\ A_1 (x_1^3 - x_1^1) + A_2 (x_2^3 - x_2^1) + A_3 (x_3^3 - x_3^1) &= 0. \end{aligned} \right\} \quad (10)$$

Let us set up a homogeneous linear system of equations with respect to the unknowns (z_1, z_2, z_3) :

$$\left. \begin{aligned} (x_1 - x_1^1) z_1 + (x_2 - x_2^1) z_2 + (x_3 - x_3^1) z_3 &= 0, \\ (x_1^2 - x_1^1) z_1 + (x_2^2 - x_2^1) z_2 + (x_3^2 - x_3^1) z_3 &= 0, \\ (x_1^3 - x_1^1) z_1 + (x_2^3 - x_2^1) z_2 + (x_3^3 - x_3^1) z_3 &= 0. \end{aligned} \right\} \quad (11)$$

Here $\mathbf{x} = (x_1, x_2, x_3)$ is an arbitrary point satisfying equation (9). By (9) and (10), system (11) is satisfied by a nontrivial vector $\mathbf{A} = (A_1, A_2, A_3)$, therefore its determinant is zero.

$$\begin{vmatrix} x_1 - x_1^1 & x_2 - x_2^1 & x_3 - x_3^1 \\ x_1^2 - x_1^1 & x_2^2 - x_2^1 & x_3^2 - x_3^1 \\ x_1^3 - x_1^1 & x_2^3 - x_2^1 & x_3^3 - x_3^1 \end{vmatrix} = 0. \quad (12)$$

We have obtained an equation of form (9), i.e. an equation of a plane which is easily verified by expanding the obtained determinant in terms of the elements of the first row. And this plane passes through the points x^1, x^2, x^3 , which follows from the properties of determinants. Our problem has been solved.

Equation (12) can also be written in the following form:

$$\begin{vmatrix} x_1 & x_2 & x_3 & 1 \\ x_1^1 & x_2^1 & x_3^1 & 1 \\ x_1^2 & x_2^2 & x_3^2 & 1 \\ x_1^3 & x_2^3 & x_3^3 & 1 \end{vmatrix} = 0. \quad (13)$$

If the second row is subtracted from the first, third, and fourth rows, then determinant (13) remains unchanged. Expanding the determinant thus obtained in terms of the elements of the fourth column, we get equation (12).

The angle between two planes. Let us specify two planes

$$\left. \begin{aligned} A_1x_1 + A_2x_2 + A_3x_3 + B &= 0, \\ A'_1x_1 + A'_2x_2 + A'_3x_3 + B' &= 0. \end{aligned} \right\} \quad (14)$$

We know that the vectors $A = (A_1, A_2, A_3)$ and $A' = (A'_1, A'_2, A'_3)$ are perpendicular to the given planes, respectively, therefore, the angle φ between A and A' is equal to the angle (dihedral) between the given planes. But the scalar product

$$AA' = |A| |A'| \cos \varphi,$$

therefore

$$\cos \varphi = \frac{AA'}{|A| |A'|} = \frac{A_1A'_1 + A_2A'_2 + A_3A'_3}{\sqrt{A_1^2 + A_2^2 + A_3^2} \sqrt{A'^2_1 + A'^2_2 + A'^2_3}}. \quad (15)$$

It is sufficient to assume that $0 \leq \varphi \leq \pi$.

Note that two intersecting planes actually form two dihedral angles φ_1 and φ_2 . Their sum is equal to π ($\varphi_1 + \varphi_2 = \pi$), and their cosines are equal to each other by absolute value but have opposite signs ($\cos \varphi_1 = -\cos \varphi_2$). If we replace, respectively, the numbers A_1, A_2, A_3 by the numbers $-A_1, -A_2, -A_3$ in equation (14), then the newly obtained equation will define the

same plane, but the angle φ in (15) will be changed for $\pi - \varphi$.

The condition for perpendicularity of planes (14) will be, obviously, $\cos \varphi = 0$, i.e.

$$\mathbf{AA}' = A_1A'_1 + A_2A'_2 + A_3A'_3 = 0. \quad (16)$$

Two planes (14) are *parallel* when, and only when, the vectors $\mathbf{A}$ and $\mathbf{A}'$ (which are perpendicular to them) are collinear, i.e. the proportionality conditions are fulfilled:

$$\frac{A_1}{A'_1} = \frac{A_2}{A'_2} = \frac{A_3}{A'_3}. \quad (17)$$

If, in addition to this, the following extended condition is fulfilled:

$$\frac{A_1}{A'_1} = \frac{A_2}{A'_2} = \frac{A_3}{A'_3} = \frac{B}{B'}, \quad (18)$$

then this means that planes (14) *coincide*, i.e. both equations (14) define one and the same plane.

Though it is impossible to divide by zero, but it is convenient to write symbolic proportions (17) or (18) with zeros in the denominators. But in this case, if, for instance, $A'_2 = 0$, then we must set $A_2 = 0$. Or, if $B' = 0$, then $B = 0$.

Examples. The equations

$$\begin{aligned} 2x + 4y + 1 &= 0, \\ x + 2y + 5 &= 0 \end{aligned}$$

determine a pair of parallel planes, and the equations

$$\begin{aligned} 2x + 4y + 2 &= 0, \\ x + 2y + 1 &= 0 \end{aligned}$$

determine a pair of coinciding planes.

The distance from a point to a plane. It is required to find the distance from the point $\mathbf{x}^0 = (x_1^0, x_2^0, x_3^0)$ to the plane L defined by the equation

$$A_1x_1 + A_2x_2 + A_3x_3 + B = 0.$$

To this effect, let us reduce the equation of L to the normal form:

$$(\mathbf{x}, \mathbf{v}) = p \quad (p \geq 0).$$

It is seen from Fig. 22 that the difference $\mathbf{x} - \mathbf{x}^0$ between the radius vector of an arbitrary point $\mathbf{x}$ belonging to the plane L and the radius vector of the point $\mathbf{x}^0$ is a vector such that the absolute value of its projection on $\mathbf{v}$ is equal to the required distance d from $\mathbf{x}^0$ to L :

$$d = |(\mathbf{x} - \mathbf{x}^0, \mathbf{v})|,$$

but

$$(\mathbf{x} - \mathbf{x}^0, \mathbf{v}) = (\mathbf{x}, \mathbf{v}) - (\mathbf{x}^0, \mathbf{v}) = p - (\mathbf{x}^0, \mathbf{v}).$$

Consequently,

$$d = |(\mathbf{x}^0, \mathbf{v}) - p|.$$

We see that, in order to compute the distance d from the point $\mathbf{x}^0$ to the plane L , we have to write the equation of L

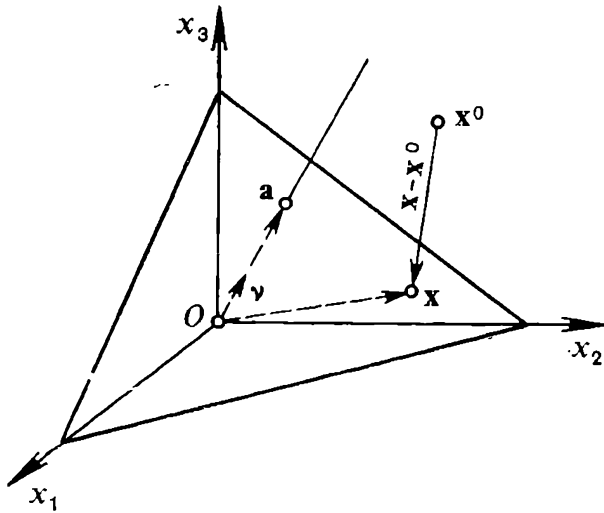


Fig. 22.

in the normal form, transpose p to the left-hand member, and substitute $\mathbf{x}^0$ for $\mathbf{x}$. The absolute value of the obtained expression is just the required number d .

It is obvious that in terms of the parameters of the plane,

$$d = \frac{|A_1 x_1^0 + A_2 x_2^0 + A_3 x_3^0 + B|}{\sqrt{A_1^2 + A_2^2 + A_3^2}}.$$

It is easy to see that if the point $\mathbf{x}^0$ and the origin are situated on different sides of the plane L (as in Fig. 22), then the vector $\mathbf{x} - \mathbf{x}^0$ forms an obtuse angle with $\mathbf{v}$ and, therefore,

$$d = -(\mathbf{x} - \mathbf{x}^0, \mathbf{v}) = (\mathbf{x}^0, \mathbf{v}) - p > 0.$$

But if the point $\mathbf{x}^0$ and the origin are found on one side of L , then an acute angle is made, and

$$d = (\mathbf{x} - \mathbf{x}^0, \mathbf{v}) = p - (\mathbf{x}^0, \mathbf{v}) > 0.$$

Consequently, in the first case $(\mathbf{x}^0, \mathbf{v}) > p$, and in the second $(\mathbf{x}^0, \mathbf{v}) < p$.

Example 1. The distance d from the point $(1, 1, 1)$ to the plane L

$$x + 2y + z = 3$$

is equal to

$$d = \left| \frac{x + 2y + z - 3}{\sqrt{1 + 4 + 1}} \right|_{x=1, y=1, z=1} = \frac{1}{\sqrt{6}}.$$

In this case the point $(1, 1, 1)$ and the origin are situated on different sides of the plane L since $M > 0$ and

$$[x + 2y + z - 3]_{x=1, y=1, z=1} = 1 > 0.$$

Problems

1. Reduce the following equations of planes

$$x - y + z = 2, \quad 2x - y + \sqrt{20}z = 10$$

to the normal form.

2. Find the angle between the two planes:

$$(a) \quad x + y + z = 1 \quad \text{and} \quad y = 0;$$

$$(b) \quad x - y + z = 2 \quad \text{and} \quad x + y + z = 3.$$

3. Write the equation of the plane passing through the points $(0, 0, 1)$, $(1, 0, 1)$, $(1, 1, 0)$.

4. Write the equation of a spherical surface with centre at the origin, tangent to the plane $2x - 6y + 7z = 2$.

Sec. 10. A Straight Line in Space

Consider in space an arbitrary straight line L . Mark on it a point $\mathbf{x}^0$ and a vector $\mathbf{a} = (a_1, a_2, a_3) \neq 0$, applied at the point $\mathbf{x}^0$ (Fig. 23). An arbitrary (moving) point of

the line L will be denoted by $\mathbf{x}$. The vector $\mathbf{x} - \mathbf{x}^0$ can be written in the form $\mathbf{x} - \mathbf{x}^0 = t\mathbf{a}$, where t is a certain number (scalar). If the real variable t runs through the

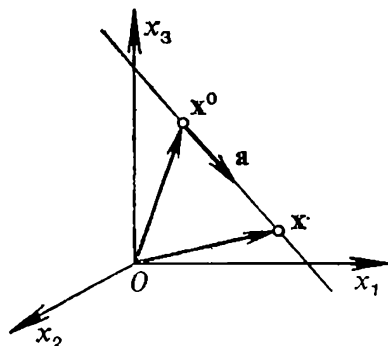


Fig. 23.

interval $(-\infty, \infty)$, then $\mathbf{x} = \mathbf{x}^0 + t\mathbf{a}$ traverses the entire length of L . Therefore, the equality

$$\mathbf{x} - \mathbf{x}^0 = t\mathbf{a} \quad (-\infty < t < \infty) \quad (1)$$

is said to be a *vector equation of the straight line passing through the point $\mathbf{x}^0$ in the same direction as $\mathbf{a}$.*

When written in terms of the coordinates, equation (1) decomposes into three equations:

$$\left. \begin{aligned} x_1 - x_1^0 &= ta_1, \\ x_2 - x_2^0 &= ta_2, \\ x_3 - x_3^0 &= ta_3. \end{aligned} \right\} \quad (1')$$

Eliminating the parameter t , we get the equations of a straight line (a system of two equations)

$$\frac{x_1 - x_1^0}{a_1} = \frac{x_2 - x_2^0}{a_2} = \frac{x_3 - x_3^0}{a_3}. \quad (1'')$$

where the numbers a_1, a_2, a_3 are not all zero. Equations (1'') are called the *equations of a straight line in the canonical form*.

Remark. It may happen that one or two of the numbers a_1, a_2, a_3 are equal to zero. Then it is still customary to write equations (1'') with a zero or even two zeros in the

denominators. Such notation then becomes symbolic, but it is convenient.

Example 1. The equations

$$\frac{x-1}{1} = \frac{y-2}{0} = \frac{z-3}{2} \quad (2)$$

define a straight line in space passing through the point (1, 2, 3) in the direction of the vector (1, 0, 2).

These equations can be replaced by the following equivalent ones:

$$y - 2 = 0 \cdot (x - 1), \quad 2(x - 1) = z - 3,$$

i.e.

$$y = 2, \quad z = 2x + 1. \quad (2')$$

Thus, the straight line under consideration is an intersection of two planes defined by equations (2').

Example 2. The equations of a straight line

$$\frac{x-1}{1} = \frac{y-2}{0} = \frac{z-3}{0}$$

are equivalent to the following ones:

$$y - 2 = 0, \quad z - 3 = 0.$$

Let there be given equations of two planes

$$A_1x + A_2y + A_3z + B = 0, \quad (3)$$

$$A'_1x + A'_2y + A'_3z + B' = 0. \quad (4)$$

If the coefficients A_j of the first of them are respectively proportional to the coefficients A'_j of the second ($A_1 : A_2 : A_3 = A'_1 : A'_2 : A'_3$), the planes (3) and (4) are parallel or even coincide (provided that $A_1 : A_2 : A_3 : B = A'_1 : A'_2 : A'_3 : B'$) (see (17) and (18) from the preceding section). Otherwise, (3) and (4) intersect along a straight line. In this case one of the determinants

$$\begin{vmatrix} A_1 & A_2 \\ A'_1 & A'_2 \end{vmatrix}, \quad \begin{vmatrix} A_1 & A_3 \\ A'_1 & A'_3 \end{vmatrix}, \quad \begin{vmatrix} A_2 & A_3 \\ A'_2 & A'_3 \end{vmatrix}$$

is not equal to zero. For the sake of definiteness, let us assume the first one:

$$\begin{vmatrix} A_1 & A_2 \\ A'_1 & A'_2 \end{vmatrix} \neq 0.$$

Then equations (3), (4) can be solved with respect to x and y , and we obtain

$$\left. \begin{aligned} x &= \alpha z + \mu, \\ y &= \beta z + \nu, \end{aligned} \right\} \quad (5)$$

where α, β, μ, ν are certain numbers. Equations (5) are equivalent to the following ones:

$$\frac{x - \mu}{\alpha} = \frac{y - \nu}{\beta} = \frac{z}{1}. \quad (6)$$

We see that the straight line defined by equations (3) and (4) passes through the point $(\mu, \nu, 0)$ in the direction of the vector $(\alpha, \beta, 1)$. The numbers α, β or one of them may be equal to zero, then equations (6) will have a symbolic character.

Example 3. The straight line defined by the equations $x = 0, y = 0$ is, obviously, the z -axis. This result can also be obtained formally. We have

$$x = 0 \cdot z, \quad y = 0 \cdot z,$$

whence

$$\frac{x}{0} = \frac{y}{0} = \frac{z}{1},$$

i.e. we have obtained the equations of the straight line passing through the origin $(0, 0, 0)$ in the direction of the vector $(0, 0, 1)$. It is clear that this line is just the z -axis.

Example 4. Find the angle between the straight lines

$$\frac{x_1 - x_1^1}{a_1} = \frac{x_2 - x_2^1}{a_2} = \frac{x_3 - x_3^1}{a_3}, \quad (7)$$

$$\frac{x_1 - x_1^2}{b_1} = \frac{x_2 - x_2^2}{b_2} = \frac{x_3 - x_3^2}{b_3}. \quad (8)$$

The vectors $\mathbf{a} = (a_1, a_2, a_3)$, $\mathbf{b} = (b_1, b_2, b_3)$ lie on these lines and, according to our agreement, they are applied

at the points $\mathbf{x}^1 = (x_1^1, x_2^1, x_3^1)$, $\mathbf{x}^2 = (x_1^2, x_2^2, x_3^2)$. Therefore, the angle φ between these vectors will be just the angle between straight lines (7) and (8):

$$\cos \varphi = \frac{a_1 b_1 + a_2 b_2 + a_3 b_3}{|\mathbf{a}| |\mathbf{b}|}.$$

Problems

1. Write the equation of the straight line passing through the point $(2, -1, 0)$ perpendicular to the plane $2x + z - 4y = 7$.
2. Write the equation of the plane passing through the point $(1, -1, 2)$ perpendicular to the straight line defined by the equations $2x + 3y = 1$, $3x - z = 1$.

Sec. 11. Orientation of Rectangular Coordinate Systems

Figures 24 and 25 represent two different coordinate systems (x_1, x_2) . They are said to be *oriented in an opposite way*. In the case presented in Fig. 24 it is possible to bring to coincidence the directions of the axes x_1 and x_2

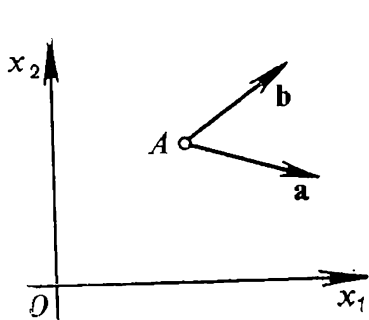


Fig. 24.

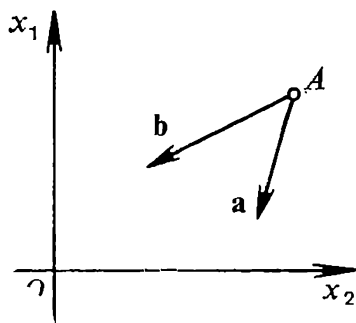


Fig. 25.

by turning the axis x_1 about the point O through an angle of $\pi/2$ anticlockwise. In the case of Fig. 25 this can be achieved only by rotating the axis x_1 clockwise. It is impossible to bring to coincidence the coordinate system represented in Fig. 24 with that depicted in Fig. 25 (so that the directions of their corresponding axes coincide) by displacing it in the plane under consideration (I) as a solid.

Figure 24, as well as Fig. 25, shows a pair of noncollinear vectors $\mathbf{a}$ and $\mathbf{b}$ emanating from a certain point A . Displacing this pair as a solid in the plane, let us bring the point A to a position in which it coincides with the origin O . Let us set ourselves the task of bringing the vector $\mathbf{a}$ into the direction of the axis x_1 and aligning the vector $\mathbf{b}$ with the axis x_2 by rotating either of the mentioned vectors about the point O . It is required that during this process the vectors $\mathbf{a}$ and $\mathbf{b}$ be situated all the time in the

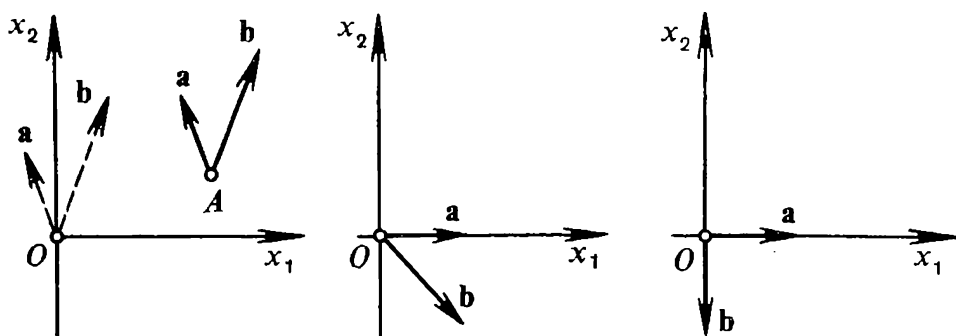


Fig. 26.

plane under consideration and that the angle between them be not equal to 0 or π . Obviously, this can always be achieved. First we rotate the system of vectors $\mathbf{a}$ and $\mathbf{b}$ as a solid about the point O to bring the vector $\mathbf{a}$ into coincidence with the positive direction of the axis x_1 . Since the vectors $\mathbf{a}$ and $\mathbf{b}$ are not collinear, the vector $\mathbf{b}$ will be found either in the upper, or lower half-plane. Then we turn the vector $\mathbf{b}$ through an angle sufficient for $\mathbf{b}$ to be brought to coincidence with the axis x_2 , $\mathbf{b}$ being not allowed to get onto the axis x_1 . Therefore, it may happen that the direction of the vector $\mathbf{b}$ will coincide with that of the axis x_2 (it is quite possible when $\mathbf{b}$ was situated in the upper half-plane) or the direction of $\mathbf{b}$ will turn out to be the same as the negative direction of the axis x_2 (see Fig. 26 illustrating this process). In the first case the pair of vectors $(\mathbf{a}, \mathbf{b})$ is said to be oriented in the same way as the coordinate system (x_1, x_2) , while in the second case the pair $(\mathbf{a}, \mathbf{b})$ is oriented in an opposite way.

The rectangular coordinate systems x_1, x_2, x_3 in space represented in Figs. 27 and 28 are also different. Considering the coordinate system shown in Fig. 27 as a solid, we come to a conclusion that, displacing it in an appropriate way, it is possible to bring to coincidence the axes x_1 and x_2 of both coordinate systems. But the positive

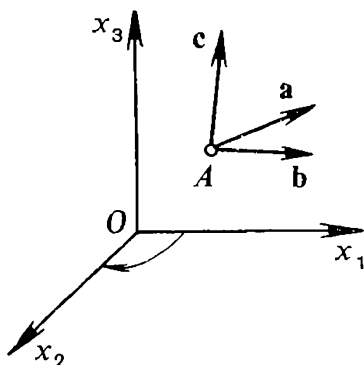


Fig. 27.

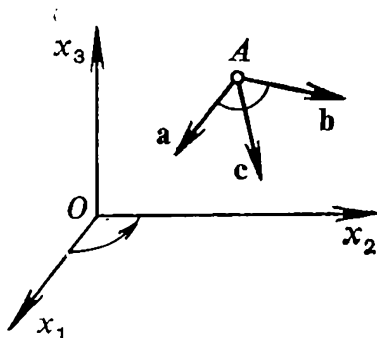


Fig. 28.

direction of the axis x_3 of the first system does not coincide with that of the second system. The systems depicted in Figs. 27 and 28 are said to be oriented in an opposite way. The system of Fig. 27 is called the *left-handed coordinate system* and that of Fig. 28 the *right-handed coordinate system*. If a screw with a right-hand (left-hand) screw thread is driven in the direction of the axis x_3 , being turned in the direction indicated by the arrow in Fig. 28 (Fig. 27), then it will perform translational motion in this direction. We can also distinguish a coordinate system according to the following rule. If a viewer located at a certain point of the positive semiaxis Ox_3 looks at the positive semiaxis Ox_2 , then the positive semiaxis Ox_1 may be found directed either leftward, or rightward. In the former case the coordinate system is called left-handed (Fig. 27), while in the latter case it is called right-handed (Fig. 28).

Vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$ are called *coplanar* if they lie in one plane or are found in parallel planes.

Let us take a system of noncoplanar vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$ applied at a certain point A and rotate the vector $\mathbf{b}$

about the point A in the plane of the vectors $\mathbf{a}$ and $\mathbf{b}$, until it turns out to be perpendicular to $\mathbf{a}$. Make sure that during this rotation the angle between $\mathbf{a}$ and $\mathbf{b}$ is never equal to 0 or π . Then we shall rotate the vector $\mathbf{c}$ about A to make it perpendicular to the vectors $\mathbf{a}$ and $\mathbf{b}$. Pay attention that the vector $\mathbf{c}$ never coincides with the plane of the vectors $\mathbf{a}$ and $\mathbf{b}$. As a result, the vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ will turn out to be mutually perpendicular. Let us now transfer this triple of vectors, as a solid, to the point O and rotate it about this point in order to make the vectors $\mathbf{a}$ and $\mathbf{b}$ coincide with the directions of the axes x_1 and x_2 , respectively. Two cases are possible here: (1) the vector $\mathbf{c}$ will be in the same direction as the axis x_3 ; (2) it will be in the opposite direction. In the first case the system of vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ is said to be *oriented in the same way as the coordinate system* x_1, x_2, x_3 , while in the second case it is said to be *oriented in an opposite way* (see Figs. 27 and 28, respectively).

Sec. 12. The Vector Product of Two Vectors

Let us specify in some rectangular system two vectors

$$\mathbf{a} = (a_1, a_2, a_3), \quad \mathbf{b} = (b_1, b_2, b_3).$$

The vector

$$\begin{aligned} \mathbf{c} = \mathbf{a} \times \mathbf{b} = [\mathbf{a} \times \mathbf{b}] = (a_2 b_3 - a_3 b_2) \mathbf{i} \\ + (a_3 b_1 - a_1 b_3) \mathbf{j} + (a_1 b_2 - a_2 b_1) \mathbf{k} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}. \end{aligned} \quad (1)$$

is then called the *vector product* of $\mathbf{a}$ and $\mathbf{b}$.

The last term of the above written chain is a "generalized determinant" whose first row consists of the vectors $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ (basis vectors of the coordinate system). The second "link" of the chain shows how to understand this generalized determinant (we expand the determinant of the third order in terms of the elements of the first row so as if $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ were numbers).

It is obvious that $[\mathbf{a} \times (-\mathbf{b})] = -[\mathbf{a} \times \mathbf{b}]$.

It is also possible to define the vector product of two vectors $\mathbf{a}$ and $\mathbf{b}$ in the following manner:

(1) if $\mathbf{a}$ and $\mathbf{b}$ are collinear, then their vector product is equal to zero;

(2) if $\mathbf{a}$ and $\mathbf{b}$ are noncollinear, then the vector $\mathbf{c}$ is perpendicular to $\mathbf{a}$ and $\mathbf{b}$ so that the system $(\mathbf{a}, \mathbf{b}, \mathbf{c})$ turns out

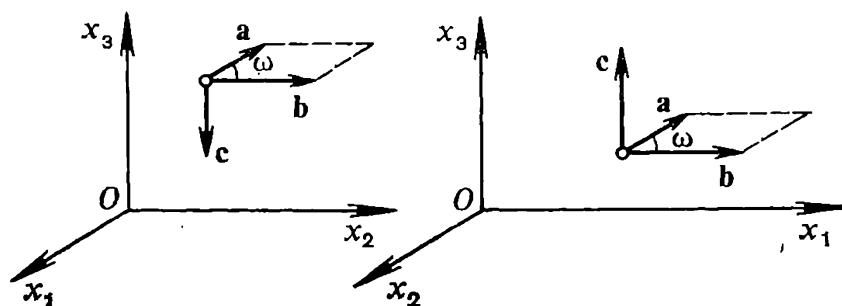


Fig. 29.

to be oriented in the same way as the given coordinate system. The length of the vector $\mathbf{c}$ is equal to

$$|\mathbf{c}| = |\mathbf{a}| \cdot |\mathbf{b}| \sin \omega \quad (0 \leq \omega \leq \pi), \quad (2)$$

where ω is the angle between $\mathbf{a}$ and $\mathbf{b}$, i.e. the length of $\mathbf{c}$ is equal to the area of the parallelogram constructed on the vectors $\mathbf{a}$ and $\mathbf{b}$ (Fig. 29).

Let us prove that the two above formulated definitions are equivalent.

If the vector $\mathbf{c} = \mathbf{0}$, then it follows from (1) that the components of the vectors $\mathbf{a}$ and $\mathbf{b}$ are proportional

$$a_1 : a_2 : a_3 = b_1 : b_2 : b_3,$$

i.e. $\mathbf{a}$ and $\mathbf{b}$ are collinear, but then the vector product is equal to zero by the second definition as well. Conversely, if $\mathbf{a}$ and $\mathbf{b}$ are collinear, then, by the second definition, $\mathbf{a} \times \mathbf{b} = \mathbf{0}$. Since the components of the vectors $\mathbf{a}$ and $\mathbf{b}$ are proportional here, then, according to the first definition, $\mathbf{c} = \mathbf{0}$.

Let now $\mathbf{a}$ and $\mathbf{b}$ be noncollinear vectors and $\mathbf{c}$ their vector product according to (1). It is obvious that

$$\begin{aligned} (\mathbf{c}, \mathbf{a}) &= (a_2 b_3 - a_3 b_2) a_1 + (a_3 b_1 - a_1 b_3) a_2 \\ &\quad + (a_1 b_2 - a_2 b_1) a_3 = 0 \end{aligned}$$

and, similarly,

$$(c, b) = 0.$$

Hence, the vector c is perpendicular to a and b .

Let us prove that the system of vectors (a, b, c) is oriented in the same way as the coordinate system (x_1, x_2, x_3) . Let us continuously rotate the vectors a and b about the point O , each time computing by them the vector c , all the time keeping a and b noncollinear. But then the vector c all the time will be not equal to zero ($c \neq 0$) and the system (a, b, c) will be noncoplanar all the time. Let us rotate the vectors a and b so as to align them with the directions of the axes x_1 and x_2 , respectively, i.e. so as to bring them to the form $a = (|a|, 0, 0)$, $b = (0, |b|, 0)$. This can always be achieved, since in this case the plane containing the vectors a and b can rotate in space. But then the vector c computed by formula (1) has the form $c = (0, 0, |a| |b|)$. We see that at the last instant of our process the vectors (a, b, c) are oriented in the same way as the axes (x_1, x_2, x_3) . But then, according to the definition of orientation (see Sec. 11), the initial system (a, b, c) is also oriented in the same way as the coordinate system (x_1, x_2, x_3) .

Thus, the vector product $c = a \times b$, defined by formula (1), is a vector perpendicular to the vectors a and b , and the system (a, b, c) is oriented in the same way as the considered coordinate system (x_1, x_2, x_3) .

Now we are going to prove formula (2). Let the vectors a and b form with the coordinate axes angles respectively equal to $\alpha_1, \alpha_2, \alpha_3$ and $\beta_1, \beta_2, \beta_3$. Since

$$a_j = |a| \cos \alpha_j, \quad b_j = |b| \cos \beta_j \quad (j = 1, 2, 3),$$

then

$$\begin{aligned} |c|^2 &= (a_2 b_3 - a_3 b_2)^2 + (a_3 b_1 - a_1 b_3)^2 \\ &+ (a_1 b_2 - a_2 b_1)^2 = |a|^2 |b|^2 [(\cos \alpha_2 \cos \beta_3 \\ &- \cos \alpha_3 \cos \beta_2)^2 \\ &+ (\cos \alpha_3 \cos \beta_1 - \cos \alpha_1 \cos \beta_3)^2 \\ &+ (\cos \alpha_1 \cos \beta_2 - \cos \alpha_2 \cos \beta_1)^2] \end{aligned}$$

$$\begin{aligned}
&= |\mathbf{a}|^2 |\mathbf{b}|^2 [(\cos^2 \alpha_1 + \cos^2 \alpha_2 + \cos^2 \alpha_3) \\
&\times (\cos^2 \beta_1 + \cos^2 \beta_2 + \cos^2 \beta_3) \\
&\quad - (\cos \alpha_1 \cos \beta_1 + \cos \alpha_2 \cos \beta_2 + \cos \alpha_3 \cos \beta_3)^2] \\
&= |\mathbf{a}|^2 |\mathbf{b}|^2 [1 - \cos^2 \omega] = |\mathbf{a}|^2 |\mathbf{b}|^2 \sin^2 \omega,
\end{aligned}$$

where ω is the angle between the vectors $\mathbf{a}$ and $\mathbf{b}$

$$(\cos \omega = \sum_{j=1}^3 \cos \alpha_j \cos \beta_j).$$

Thus, we have proved (2).

Hence, we completely proved that from the definition of the vector product by formula (1) there follows its second definition.

Conversely, if a vector obeys the second definition, then it is unique, since there can be only one vector perpendicular to $\mathbf{a}$ and $\mathbf{b}$ whose length is equal to the area of the

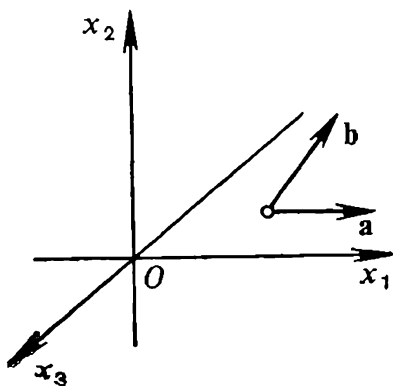


Fig. 30.

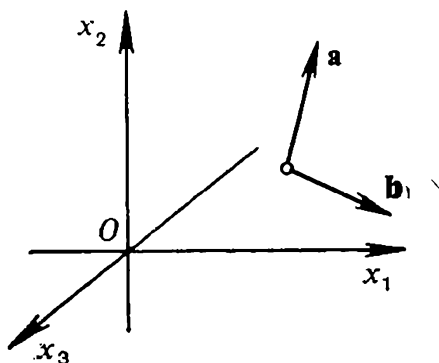


Fig. 31.

parallelogram constructed on $\mathbf{a}$, $\mathbf{b}$, such that the system $(\mathbf{a}, \mathbf{b}, \mathbf{c})$ is oriented in the same way as the system (x_1, x_2, x_3) . But then it is the vector $\mathbf{c}$ defined by formula (1), since the latter, as we saw, possesses the mentioned properties.

We should like to note once again that the condition $\mathbf{a} \times \mathbf{b} = \mathbf{0}$ is a *necessary and sufficient condition for the collinearity of the vectors $\mathbf{a}$ and $\mathbf{b}$* .

Now consider two nonzero plane vectors

$$\mathbf{a} = (a_1, a_2), \quad \mathbf{b} = (b_1, b_2) \quad (3)$$

in a rectangular coordinate system (x_1, x_2) (Figs. 30, 31). Regarding the plane under consideration as given in space, let us add to the axes x_1, x_2 an axis x_3 which is perpendicular to them. The vectors $\mathbf{a}, \mathbf{b}$ will then have the following coordinates:

$$\mathbf{a} = (a_1, a_2, 0), \quad \mathbf{b} = (b_1, b_2, 0).$$

Their vector product is equal to

$$\mathbf{c} = \mathbf{a} \times \mathbf{b} = \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} \mathbf{k}, \quad (4)$$

where $\mathbf{k}$ is a basis vector of the axis x_3 . According to the definition of the vector product, the system $(\mathbf{a}, \mathbf{b}, \mathbf{c})$ is oriented in the same way as the coordinate system (x_1, x_2, x_3) . Therefore, obviously, if the determinant

$$\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} > 0,$$

then the system of vectors $(\mathbf{a}, \mathbf{b})$ must be oriented in the way the coordinate axes (x_1, x_2) are oriented. And if the determinant

$$\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} < 0,$$

then the system $(\mathbf{a}, \mathbf{b})$ has an orientation opposite to that of (x_1, x_2) . The first arrangement of the vectors $(\mathbf{a}, \mathbf{b})$ is shown in Fig. 30, the second in Fig. 31. We also know that the area of the parallelogram constructed on the vectors $\mathbf{a}$ and $\mathbf{b}$ is equal to (see (4))

$$S = |\mathbf{c}| = \left| \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} \right|,$$

i.e. to the absolute value of the determinant

$$\Delta = \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}. \quad (5)$$

Thus, we have proved that: (1) the absolute value of determinant (5) is equal to the area of the parallelogram

constructed on the vectors $\mathbf{a} = (a_1, a_2)$ and $\mathbf{b} = (b_1, b_2)$; (2) the sign of determinant (5) depends on the arrangement of these vectors relative to the rectangular coordinate system (x_1, x_2) . To the plus sign there corresponds a system $(\mathbf{a}, \mathbf{b})$ oriented the same as (x_1, x_2) , and to the minus sign there corresponds a system $(\mathbf{a}, \mathbf{b})$ oriented in an opposite manner.

The following properties hold true:

$$\mathbf{a} \times \mathbf{b} = -[\mathbf{b} \times \mathbf{a}], \quad (6)$$

$$\mathbf{a} \times (\alpha \mathbf{b}) = \alpha [\mathbf{a} \times \mathbf{b}], \quad (7)$$

$$\mathbf{a} \times (\mathbf{b} + \mathbf{c}) = [\mathbf{a} \times \mathbf{b}] + [\mathbf{a} \times \mathbf{c}], \quad (8)$$

where $\mathbf{a}, \mathbf{b}, \mathbf{c}$ are arbitrary vectors, and α is a scalar.

Expressing the vector products entering into equations (6), (7), (8) by formula (1) in terms of the components of the vectors

$$\mathbf{a} = (a_1, a_2, a_3), \quad \mathbf{b} = (b_1, b_2, b_3), \quad \mathbf{c} = (c_1, c_2, c_3),$$

we easily obtain these equalities.

Formulas (6) and (7) readily follow from geometrical considerations as well. Let $\mathbf{a}$ and $\mathbf{b}$ be noncollinear vectors. If $\mathbf{a}$ and $\mathbf{b}$ are interchanged in the vector product, then both the area of the parallelogram constructed on $\mathbf{a}$ and $\mathbf{b}$, and the perpendicular to $\mathbf{a}$ and $\mathbf{b}$ will remain unchanged. Only the direction of $\mathbf{c} = \mathbf{a} \times \mathbf{b}$ will reverse thus changing the orientation of $(\mathbf{a}, \mathbf{b}, \mathbf{a} \times \mathbf{b})$.

The multiplication of the vector $\mathbf{b}$ by a positive number α enlarges α times only the area of the parallelogram constructed on $\mathbf{a}$ and $\mathbf{b}$, the direction of the vector product remaining unaltered. But if $\alpha < 0$, then

$$\begin{aligned} \mathbf{a} \times (\alpha \mathbf{b}) &= \mathbf{a} \times (-|\alpha| \mathbf{b}) = |\alpha| [\mathbf{a} \times (-\mathbf{b})] \\ &= -|\alpha| [\mathbf{a} \times \mathbf{b}] = \alpha [\mathbf{a} \times \mathbf{b}]. \end{aligned}$$

One more remark: it also follows from (6) and (7) that $(\alpha \mathbf{a}) \times \mathbf{b} = -[\mathbf{b} \times (\alpha \mathbf{a})] = -\alpha [\mathbf{b} \times \mathbf{a}] = \alpha [\mathbf{a} \times \mathbf{b}]$.

Example 1. Determine the angle A of the triangle ABC with the vertices: $A = (1, 2, 3)$, $B = (2, 2, 2)$, $C = (1, 2, 4)$.

Let us denote the required angle by ω , which is actually the angle between the vectors $\vec{AB}$ and $\vec{AC}$. By the second definition of the vector product, we have

$$\sin \omega = \frac{|\vec{AB} \times \vec{AC}|}{|\vec{AB}| |\vec{AC}|},$$

where $\vec{AB} = (1, 0, -1)$, $\vec{AC} = (0, 0, 1)$; $|\vec{AB}| = \sqrt{2}$, $|\vec{AC}| = 1$,

$$\vec{AB} \times \vec{AC} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 0 & -1 \\ 0 & 0 & 1 \end{vmatrix} = -\mathbf{j}.$$

Hence

$$\sin \omega = \frac{1}{\sqrt{2}}, \quad \omega = \frac{\pi}{4}, \quad \omega = \frac{3\pi}{4}.$$

Since $\vec{BC} = (-1, 0, 2)$ and $|\vec{BC}|^2 = 5 > |\vec{AC}|^2 + |\vec{AB}|^2 = 1 + 2 = 3$, it is necessary to take $\omega = 3\pi/4$.

Remark. If in the triangle ABC the angle A is a right one, then $|\vec{BC}|^2 = BC^2 = AC^2 + AB^2$; if A is obtuse, then $BC^2 > AC^2 + AB^2$; if A is acute, then $BC^2 < AC^2 + AB^2$.

Sec. 13. Triple Scalar Product

The *triple scalar product of three vectors* $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ is a scalar equal to the scalar product of the vector $\mathbf{a} \times \mathbf{b}$ by the vector $\mathbf{c}$:

$$\begin{aligned} (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} &= (a_2 b_3 - a_3 b_2) c_1 + (a_3 b_1 - a_1 b_3) c_2 \\ &+ (a_1 b_2 - a_2 b_1) c_3 = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}. \end{aligned} \quad (1)$$

By virtue of the definition of a scalar product

$$(\mathbf{a} \times \mathbf{b}) \mathbf{c} = |\mathbf{a} \times \mathbf{b}| \operatorname{pr}_{\mathbf{a} \times \mathbf{b}} \mathbf{c} = (|\mathbf{a}| \cdot |\mathbf{b}| \sin \omega) \operatorname{pr}_{\mathbf{a} \times \mathbf{b}} \mathbf{c}.$$

Therefore we may, obviously, assert that the triple product $(\mathbf{a} \times \mathbf{b}) \mathbf{c}$ is equal to the volume of the parallelepiped constructed on the vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$. The latter is taken with the sign $+$ or $-$ depending on whether the system of vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ is oriented as the coordinate system x_1, x_2, x_3 or in an opposite way. Note that $|\operatorname{pr}_{\mathbf{a} \times \mathbf{b}} \mathbf{c}|$ is equal to the altitude of the parallelepiped.

There exist the following equalities

$$(\mathbf{a} \times \mathbf{b}) \mathbf{c} = (\mathbf{c} \times \mathbf{a}) \mathbf{b} = (\mathbf{b} \times \mathbf{c}) \mathbf{a}, \quad (2)$$

which easily follow from the properties of determinant (1).

If the vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ lie in one and the same plane, then

$$(\mathbf{a} \times \mathbf{b}) \mathbf{c} = 0,$$

since $\mathbf{a} \times \mathbf{b}$ is perpendicular to the vector $\mathbf{c}$. Conversely, if $(\mathbf{a} \times \mathbf{b}) \mathbf{c} = 0$, then the vector $\mathbf{c}$ is perpendicular to the vector $\mathbf{a} \times \mathbf{b}$ and, consequently, lies in the plane containing the vectors $\mathbf{a}$ and $\mathbf{b}$ or in a plane parallel to this plane.

Thus, the condition

$$(\mathbf{a} \times \mathbf{b}) \mathbf{c} = 0$$

is a *necessary and sufficient condition for the coplanarity of three vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$.*

Example 1. Find the condition for belonging of four points to one plane.

Let there be given four points $A_j = (x_j, y_j, z_j)$ ($j = 1, 2, 3, 4$). If these points lie in one and the same plane, then the vectors $\overrightarrow{A_1 A_2}$, $\overrightarrow{A_1 A_3}$, $\overrightarrow{A_1 A_4}$ also lie in this plane and, consequently, their triple product is zero:

$$\overrightarrow{A_1 A_2} \times \overrightarrow{A_1 A_3} \cdot \overrightarrow{A_1 A_4} = \begin{vmatrix} x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\ x_3 - x_1 & y_3 - y_1 & z_3 - z_1 \\ x_4 - x_1 & y_4 - y_1 & z_4 - z_1 \end{vmatrix} = 0.$$

This is just the condition of belonging of four points to a single plane (cf. (12) from Sec. 9).

Sec. 14. Linearly Independent System of Vectors

Let us specify in R_n a system of k vectors

$$\mathbf{a}^s = (a_{s1}, a_{s2}, \dots, a_{sn}) \quad (s = 1, \dots, k). \quad (1)$$

System (1) is defined as *linearly independent* if from the vector equation

$$\lambda_1 \mathbf{a}^1 + \lambda_2 \mathbf{a}^2 + \dots + \lambda_k \mathbf{a}^k = \mathbf{0}, \quad (2)$$

where $\lambda_1, \lambda_2, \dots, \lambda_k$ are numbers, it follows that

$$\lambda_1 = \lambda_2 = \dots = \lambda_k = 0.$$

System (1) is called *linearly dependent* if there exist numbers $\lambda_1, \lambda_2, \dots, \lambda_k$ not all equal to zero, for which equation (2) is fulfilled. If, for the sake of definiteness, we assume that $\lambda_k \neq 0$, then it follows from (2) that

$$\mathbf{a}^k = \mu_1 \mathbf{a}^1 + \dots + \mu_{k-1} \mathbf{a}^{k-1}, \quad (3)$$

where

$$\mu_1 = -\frac{\lambda_1}{\lambda_k}, \dots, \mu_{k-1} = -\frac{\lambda_{k-1}}{\lambda_k}.$$

Thus, if a system of k vectors is linearly dependent, then one of them is said to be a *linear combination* of the remaining vectors, or, in other words, one of them *depends on the remaining vectors*.

Since all the time we shall speak of a *linear* dependence, the term *linear* will sometimes be omitted. We shall also speak of *dependent* and *independent vectors* instead of *dependent* or *independent system of vectors*.

One vector $\mathbf{a}^1$ also forms a system which is linearly independent if $\mathbf{a}^1 \neq \mathbf{0}$ and dependent if $\mathbf{a}^1 = \mathbf{0}$.

If an entire system of vectors $\mathbf{a}^1, \dots, \mathbf{a}^k$ is linearly independent, then any part of this system $\mathbf{a}^1, \dots, \mathbf{a}^s$ ($s < k$) is linearly independent all the more. Otherwise, we could find a nontrivial system of numbers $\lambda_1, \dots, \lambda_s$ for which

$$\lambda_1 \mathbf{a}^1 + \dots + \lambda_s \mathbf{a}^s = \mathbf{0},$$

would be fulfilled. But then for the system $\lambda_1, \dots, \lambda_s$,
 $\underbrace{\quad}_{k-s \text{ times}}$

$0, \dots, 0$ which is also nontrivial, the following would

hold true

$$\lambda_1 \mathbf{a}^1 + \dots + \lambda_s \mathbf{a}^s + 0 \cdot \mathbf{a}^{s+1} + \dots + 0 \cdot \mathbf{a}^h = 0.$$

Hence it follows that if a system of vectors $\mathbf{a}^1, \dots, \mathbf{a}^s$ is linearly dependent, then any augmented system

$$\mathbf{a}^1, \dots, \mathbf{a}^s, \mathbf{a}^{s+1}, \dots, \mathbf{a}^h$$

possesses the same property. In particular, a system of vectors containing a zero vector is always linearly dependent.

Let us form a matrix determined by the vectors of system (1):

$$A = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{k1} & a_{k2} & \dots & a_{kn} \end{vmatrix}.$$

Theorem 1. *If rank $A = k$, i.e. rank A is equal to the number of vectors, then system (1) is linearly independent.*

But if rank $A < k$, then system (1) is linearly dependent.

Example 1. Two vectors $\mathbf{a}^1 = (a_{11}, a_{12})$, $\mathbf{a}^2 = (a_{21}, a_{22})$ in R_2 form a linearly independent system if the determinant

$$\Delta = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \neq 0,$$

since the vector equation

$$\lambda_1 \mathbf{a}^1 + \lambda_2 \mathbf{a}^2 = 0 \tag{4}$$

is equivalent to two equations for the corresponding components

$$\left. \begin{aligned} a_{11}\lambda_1 + a_{21}\lambda_2 &= 0, \\ a_{12}\lambda_1 + a_{22}\lambda_2 &= 0. \end{aligned} \right\} \tag{5}$$

But if $\Delta \neq 0$, then system (5) has a unique trivial solution:

$$\lambda_1 = \lambda_2 = 0. \tag{6}$$

And if $\Delta = 0$, then equations (5) are satisfied by a certain nontrivial system (λ_1, λ_2) , i.e. for $\Delta = 0$ the system of vectors $\mathbf{a}^1, \mathbf{a}^2$ is linearly dependent.

Obviously, to say that in R_2 the vectors $\mathbf{a}^1$ and $\mathbf{a}^2$ are collinear is the same as to say that they are linearly de-

pendent. But then to state that the vectors $\mathbf{a}^1$ and $\mathbf{a}^2$ are not collinear is the same as to say that they are linearly independent.

Example 2. A system of vectors $\mathbf{a}^1, \mathbf{a}^2, \dots, \mathbf{a}^k$ ($k \geq 3$) in R_2 is always linearly dependent. Geometrically this

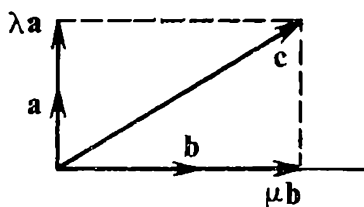


Fig. 32.

is clear from Fig. 32: if $\mathbf{c}$ is an arbitrary vector and $\mathbf{a}, \mathbf{b}$ are noncollinear vectors, then we can always indicate numbers λ, μ such that

$$\mathbf{c} = \lambda \mathbf{a} + \mu \mathbf{b}.$$

This shows that the system $\mathbf{a}, \mathbf{b}, \mathbf{c}$ is linearly dependent. But if $\mathbf{a}$ and $\mathbf{b}$ are collinear vectors, then they are linearly dependent. And $\mathbf{a}, \mathbf{b}, \mathbf{c}$ are linearly dependent all the more.

According to Theorem 1, to investigate the pair of vectors $\mathbf{a}^1, \mathbf{a}^2$, we must write the matrix formed from their coordinates

$$A = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}.$$

In the given case $k = 2$.

(a) If $\text{rank } A = 1 < 2 = k$, then the theorem states that the vectors $\mathbf{a}^1, \mathbf{a}^2$ are linearly dependent.

(b) If $\text{rank } A = 2 = k$, then the vectors $\mathbf{a}^1, \mathbf{a}^2$ are linearly independent.

This coincides with the above conclusions, since in the case of (a) $\Delta = 0$ and (b) $\Delta \neq 0$.

The fact that three arbitrary vectors $\mathbf{a}^1, \mathbf{a}^2, \mathbf{a}^3$ in R_2 are linearly dependent is also stipulated by the theorem. As a matter of fact

$$\text{rank } A \leq 2 < 3 = k.$$

Example 3. In three-dimensional space R_3 two vectors

$$\mathbf{a}^1 = (a_{11}, a_{12}, a_{13}), \mathbf{a}^2 = (a_{21}, a_{22}, a_{23})$$

are linearly dependent when, and only when, they are collinear.

Indeed, let $\mathbf{a}^1, \mathbf{a}^2$ be collinear. If one of the given vectors is zero, then they are linearly dependent. But if $\mathbf{a}^1$ and $\mathbf{a}^2$ are collinear and nonzero, then

$$\mathbf{a}^1 = \lambda \mathbf{a}^2,$$

where λ is some number. This means that $\mathbf{a}^1, \mathbf{a}^2$ are linearly dependent.

Conversely, if $\mathbf{a}^1, \mathbf{a}^2$ are linearly dependent, then one of them depends on the other, for instance,

$$\mathbf{a}^2 = \mu \mathbf{a}^1,$$

i.e. the vectors are collinear.

Consider the matrix

$$A = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{vmatrix}.$$

Since the elements of its rows are proportional, then

$$\text{rank } A = 1 < 2 = k,$$

i.e. our assertion conforms to Theorem 1.

Example 4. Let us now consider three vectors in R_3 :

$$\mathbf{a}^1 = (a_{11}, a_{12}, a_{13}),$$

$$\mathbf{a}^2 = (a_{21}, a_{22}, a_{23}),$$

$$\mathbf{a}^3 = (a_{31}, a_{32}, a_{33}).$$

Let

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}.$$

The vector equation

$$\lambda_1 \mathbf{a}^1 + \lambda_2 \mathbf{a}^2 + \lambda_3 \mathbf{a}^3 = \mathbf{0} \quad (7)$$

is equivalent to the system of three equations

$$\left. \begin{aligned} a_{11}\lambda_1 + a_{21}\lambda_2 + a_{31}\lambda_3 &= 0, \\ a_{12}\lambda_1 + a_{22}\lambda_2 + a_{32}\lambda_3 &= 0, \\ a_{13}\lambda_1 + a_{23}\lambda_2 + a_{33}\lambda_3 &= 0. \end{aligned} \right\} \quad (7')$$

If $\Delta \neq 0$, then system (7') has a unique trivial solution $\lambda_1 = \lambda_2 = \lambda_3 = 0$. But then equation (7) also has a unique trivial solution and the system of vectors ($\mathbf{a}^1$, $\mathbf{a}^2$, $\mathbf{a}^3$) is linearly independent.

If $\Delta = 0$, then system (7') and, consequently, equation (7) have a nontrivial solution (λ_1 , λ_2 , λ_3). But then the system of vectors ($\mathbf{a}^1$, $\mathbf{a}^2$, $\mathbf{a}^3$) is linearly dependent. The following peculiarities are distinguished here:

(1) Let $\text{rank } A = 1$, where

$$A = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}.$$

Then at least one of the rows of A , let, for the sake of definiteness, the first one, has at least one nonzero element. Consider the matrix

$$A' = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{vmatrix}. \quad (8)$$

It has rank 1, therefore all the second-order determinants generated by this matrix are equal to zero

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix} = \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} = 0.$$

But then, obviously the components of the vectors $\mathbf{a}^1$ and $\mathbf{a}^2$ are proportional

$$\frac{a_{21}}{a_{11}} = \frac{a_{22}}{a_{12}} = \frac{a_{23}}{a_{13}} = \lambda,$$

i.e.

$$a_{21} = a_{11}\lambda, \quad a_{22} = a_{12}\lambda, \quad a_{23} = a_{13}\lambda$$

or

$$\mathbf{a}^2 = \lambda \mathbf{a}^1.$$

Let rank $A = k$. Then, obviously, $k \leq n$. There exists a nonzero determinant of order k generated by matrix A . Without loss of generality, we may regard that this is a determinant made up from the coefficients of the first k equations of system (2'):

$$\left\{ \begin{aligned} \lambda_1 a_{11} + \dots + \lambda_k a_{k1} &= 0, \\ . &. \\ \lambda_1 a_{1k} + \dots + \lambda_k a_{kk} &= 0, \end{aligned} \right\} \quad (10)$$
$$\begin{vmatrix} a_{11} & \dots & a_{k1} \\ . & . & . \\ a_{1k} & \dots & a_{kk} \end{vmatrix} \neq 0.$$

Since homogeneous system (10) has a nonzero determinant, it has only a trivial solution

$$\lambda_1 = \lambda_2 = \dots = \lambda_k = 0,$$

and system of vectors (1) is linearly independent. The statement of the theorem has been proved for this case.

Let now $\text{rank } A = s < k$. We renumber system (2') and variables λ_j so that

$$\begin{vmatrix} a_{11} & \dots & a_{s1} \\ \vdots & & \vdots \\ a_{1s} & \dots & a_{ss} \end{vmatrix} \neq 0.$$

Then the system of the first s equations can be written as follows:

[illegible]

Setting $\lambda_{s+1} = \dots = \lambda_k = 1$, we can solve the given system with respect to $\lambda_1, \dots, \lambda_s$. As the result, we get a nontrivial solution

$$\lambda_1, \dots, \lambda_k \quad (11)$$

of the first s equations of (2'). Now any $k - s$ equations of (2') can be added to these s equations, and the newly obtained system will have the found nontrivial system as its solution all the same. In order to explain this

i.e. the i th coordinate of the vector y is written with the aid of the i th row of A . We shall write this operator in a brief notation:

$$y = Ax \quad (x, y \in R_n) \quad (2'')$$

In particular, in the case of one-dimensional space R_1 the vectors x and y are numbers, and operator $(2'')$ turns into a function:

$$y = a_{11}x.$$

Operator $(2'')$ is linear. This means that it satisfies the conditions

$$A(\alpha x + \beta x') = \alpha Ax + \beta Ax'$$

for any vectors $x, x' \in R_n$ and numbers α, β . Indeed,

$$\sum_{j=1}^n a_{ij}(\alpha x_j + \beta x'_j) = \alpha \sum_{j=1}^n a_{ij}x_j + \beta \sum_{j=1}^n a_{ij}x'_j \quad (i = 1, \dots, n).$$

If the matrices $A = \|a_{kl}\|$ and $B = \|b_{kl}\|$ are equal, i.e. have equal corresponding elements $a_{kl} = b_{kl}$, then they determine identically equal operators:

$$Ax = Bx, \quad \forall x \in R_n. \quad (3)$$

Conversely, it follows from equation (3) that

$$a_{kl} = b_{kl} \quad (k, l = 1, \dots, n),$$

i.e. the equality of the matrices A and B ($A = B$). This can be easily verified by setting in (3)

$$x = x^l = (0, \dots, 0, 1, 0, \dots, 0),$$

where 1 occupies the l th place ($l = 1, \dots, n$).

Hence, to different matrices A there correspond different operators, if two matrices A_1 and A_2 differ from each other at least by one element, then there necessarily exists a vector x for which

$$A_1x \neq A_2x.$$

Let, in addition to A , there be given another operator B defined by a square matrix of order n

$$B = \|b_{kl}\|.$$

To every $\mathbf{x} \in R_n$ there corresponds (due to operator A) a vector $\mathbf{y} \in R_n$ to which, with the aid of operator B , there corresponds a vector $\mathbf{z}$ with the components computed by the formulas

$$z_k = \sum_{i=1}^n b_{ki} y_i \quad (k = 1, \dots, n)$$

As the result, we get a composite linear operator

$$\mathbf{z} = B A \mathbf{x} \quad (\mathbf{x} \in R_n), \quad (4)$$

where

$$\begin{aligned} z_k &= \sum_{i=1}^n b_{ki} y_i = \sum_{i=1}^n b_{ki} \sum_{j=1}^n a_{ij} x_j \\ &= \sum_{j=1}^n \left(\sum_{i=1}^n b_{ki} a_{ij} \right) x_j = \sum_{j=1}^n \gamma_{kj} x_j \end{aligned}$$

with a matrix $|\gamma_{kj}|$ called the *product of matrices* B and A , and denoted as

$$BA = \|\gamma_{kj}\|, \quad (5)$$

where

$$\gamma_{kj} = \sum_{i=1}^n b_{ki} a_{ij} \quad (k, j = 1, \dots, n). \quad (6)$$

i.e. in order to obtain the element γ_{kj} of the matrix BA (belonging to its k th row and j th column), we have to multiply the elements of the k th row of matrix B by the corresponding elements of the j th column of matrix A and add the result.

The determinant of matrix BA is equal to the product of the determinants of matrices B and A :

$$|BA| = |B| |A|. \quad (7)$$

This property follows from the formula for the product of determinants (see Sec. 2, property (j)).

Let the matrix of operator A (see (1)) have a determinant which is not equal to zero:

$$\Delta = |a_{ki}| \neq 0.$$

In this case (see Sec. 4, Theorem 1) system of equations (2), or, which is the same, the operator equation $\mathbf{y} = A\mathbf{x}$ has

a unique solution $\mathbf{x} \in R_n$ for any given $\mathbf{y} \in R_n$. And the formulas by which $\mathbf{x}$ is found for a given $\mathbf{y} \in R_n$ have the form

$$x_j = \sum_{s=1}^n b_{js} y_s \quad (j=1, \dots, n). \quad (8)$$

Here

$$b_{js} = A_{sj}/\Delta \quad (s, j = 1, \dots, n) \quad (9)$$

(see Sec. 4, (3')) where A_{sj} is the cofactor of the element a_{sj} in the determinant Δ .

However, now it is of importance for us only to note that the numbers b_{js} are elements of the matrix

$$B = \| b_{js} \|,$$

possessing the following remarkable properties:

$$BA\mathbf{x} = \mathbf{x}, \quad \forall \mathbf{x} \in R_n, \quad (10)$$

$$AB\mathbf{y} = \mathbf{y}, \quad \forall \mathbf{y} \in R_n. \quad (11)$$

Indeed, by means of operator A , an arbitrary vector $\mathbf{x} \in R_n$ turns into a certain vector $\mathbf{y}$ which, by means of operator B , turns back into $\mathbf{x}$. On the other hand, to every $\mathbf{y} \in R_n$ there corresponds, via operator B (see (8)), a certain $\mathbf{x}$ such that $A\mathbf{x} = \mathbf{y}$.

In equation (11) we may, obviously, replace $\mathbf{y}$ by another letter, therefore, we have obtained the identities

$$BA\mathbf{x} = AB\mathbf{x} = \mathbf{x}, \quad \forall \mathbf{x}.$$

The operator $\mathbf{x} = E\mathbf{x}$ ($\forall \mathbf{x} \in R_n$) is called a *unit operator*. Its corresponding matrix has the form

$$E = \left\| \begin{array}{cccc} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ . & . & . & . \\ 0 & 0 & \dots & 1 \end{array} \right\|$$

and is termed a *unit matrix*. We have proved that

$$AB = BA = E.$$

Operator B , possessing this property, is called an *inverse* of operator A and is denoted by A^{-1} . Accordingly, its

matrix is also denoted by A^{-1} . The elements of matrix A^{-1} are found by the elements of matrix A with the aid of formulas (9).

We have proved that *if the determinant $|A|$ of a square matrix A is not equal to zero, then it has an inverse matrix A^{-1}* . Thus, for A^{-1} the following properties are fulfilled:

$$A^{-1}A = AA^{-1} = E.$$

If the determinant of matrix A is equal to zero ($|A| = 0$), then it has no inverse matrix. It is sufficient to say that the equation $y = Ax$ has a solution not for any y .

Note that the property $AA^{-1}y = y$, if it is fulfilled, states that to *every* $y \in R_n$ there corresponds (with the aid of operator A^{-1}) an x such that it is the solution of the equation $y = Ax$.

Remark. The operation of multiplication of matrices can be extended to nonsquare matrices B and A , provided that the number of columns of matrix B coincides with the number of rows of matrix A . Then the multiplication of matrices is carried out by formulas similar to (6). For instance, if

$$B = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix},$$

then

$$BA = \begin{pmatrix} 4 & 3 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

In this case we cannot consider the product AB , since matrix A has two columns, while matrix B has three rows.

For square matrices A and B the products AB and BA have sense, but AB is by far not always equal to BA .

For instance, if

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$$

then

$$AB = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}, \quad BA = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix},$$

i.e. $AB \neq BA$. It is easy to verify that $(AB)C = A(BC)$.

If A is a linear operator, then the notation Ax may be regarded as the product of matrix A by a one-column matrix

$$x = \begin{pmatrix} x_1 \\ \cdot \\ \cdot \\ x_n \end{pmatrix}.$$

Let there be given linear operators A and B . Their *sum* is an operator $A + B$ defined by the equation

$$(A + B)x = Ax + Bx, \quad \forall x \in R_n.$$

Obviously, the matrix of operator $A + B$ coincides with the matrix equal to the sum of the matrices of operators A and B .

It is easy to verify that

$$A(B + C) = AB + AC.$$

Example 1. Find the inverse of the matrix

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}.$$

Let us denote the elements of the inverse matrix A^{-1} by b_{js} . We have $\Delta = 1$, $b_{js} = A_{sj}$,

$$A_{11} = \begin{vmatrix} 1 & 1 \\ 1 & 0 \end{vmatrix} = -1, \quad A_{12} = - \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} = 1,$$

$$A_{13} = \begin{vmatrix} 0 & 1 \\ 1 & 1 \end{vmatrix} = -1,$$

$$A_{21} = - \begin{vmatrix} 2 & 0 \\ 1 & 0 \end{vmatrix} = 0, \quad A_{22} = \begin{vmatrix} 1 & 0 \\ 1 & 0 \end{vmatrix} = 0,$$

$$A_{23} = - \begin{vmatrix} 1 & 2 \\ 1 & 1 \end{vmatrix} = 1,$$

$$A_{31} = \begin{vmatrix} 2 & 0 \\ 1 & 1 \end{vmatrix} = 2, \quad A_{32} = - \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = -1,$$

$$A_{33} = \begin{vmatrix} 1 & 2 \\ 0 & 1 \end{vmatrix} = 1.$$

Hence,

$$\begin{aligned} b_{11} &= A_{11} = -1, & b_{12} &= A_{21} = 0, & b_{13} &= A_{31} = 2, \\ b_{21} &= A_{12} = 1, & b_{22} &= A_{22} = 0, & b_{23} &= A_{32} = -1, \\ b_{31} &= A_{13} = -1, & b_{32} &= A_{23} = 1, & b_{33} &= A_{33} = 1. \end{aligned}$$

Thus,

$$A^{-1} = \begin{pmatrix} -1 & 0 & 2 \\ 1 & 0 & -1 \\ -1 & 1 & 1 \end{pmatrix}$$

Sec. 16. Bases in R_n

In space R_n let us introduce n vectors

$$\left. \begin{aligned} \mathbf{i}^1 &= (1, 0, \dots, 0), \\ \mathbf{i}^2 &= (0, 1, 0, \dots, 0), \\ &\dots \dots \dots \dots \dots \dots \dots \\ \mathbf{i}^n &= (0, 0, \dots, 0, 1), \end{aligned} \right\} \quad (1)$$

called *basis vectors of the axes of the R_n space*.

The *axis x_k* of the R_n space is defined as the set of points of the form $(0, \dots, 0, x_k, 0, \dots, 0)$ where x_k occupies the k th place and runs all real values, and the vector $\mathbf{i}^k$ is called the *basis vector of the axis x_k* .

If $\mathbf{a} = (x_1, \dots, x_n)$ is an arbitrary vector in R_n , then it can, obviously, be written in the form of a linear combination of vectors (1) as follows:

$$\mathbf{a} = x_1 \mathbf{i}^1 + x_2 \mathbf{i}^2 + \dots + x_n \mathbf{i}^n. \quad (2)$$

Let us denote the matrix of vectors $(\mathbf{a}^1, \dots, \mathbf{a}^n)$ by

$$A = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}.$$

The transition from the basis $(\mathbf{i}^1, \dots, \mathbf{i}^n)$ to basis $(\mathbf{a}^1, \dots, \mathbf{a}^n)$ is realized with the aid of matrix A :

$$\mathbf{a}^k = \sum_{s=1}^n a_{ks} \mathbf{i}^s \quad (k=1, \dots, n), \quad (7)$$

i.e. the vector $\mathbf{a}^k$ is expressed in terms of vectors $\mathbf{i}^s$ with the aid of the k th row of matrix A . The transition from $(\mathbf{a}^1, \dots, \mathbf{a}^n)$ to $(\mathbf{i}^1, \dots, \mathbf{i}^n)$ is effected by means of inverse matrix A^{-1} (see (9) from the preceding section)

$$\mathbf{i}^s = \sum_{l=1}^n b_{sl} \mathbf{a}^l \quad (s=1, \dots, n), \quad (8)$$

whose elements are computed by the formulas $b_{sl} = \frac{A_{ls}}{\Delta}$, where A_{ls} is the cofactor of the element a_{ls} in the determinant Δ (note that the element b_{sl} belonging to the s th row and l th column is expressed in terms of the cofactor A_{ls} of the element a_{ls} belonging to the l th row and s th column). Note also that

$$\mathbf{a} = \sum_{l=1}^n x'_l \mathbf{a}^l = \sum_{s=1}^n x_s \mathbf{i}^s = \sum_{s=1}^n x_s \sum_{l=1}^n b_{sl} \mathbf{a}^l = \sum_{l=1}^n \left(\sum_{s=1}^n b_{sl} x_s \right) \mathbf{a}^l,$$

whence

$$x'_l = \sum_{s=1}^n b_{sl} x_s, \quad (9)$$

i.e. the transition from the coordinates $(x_1, \dots, x_n)$ to $(x'_1, \dots, x'_n)$ is effected with the aid of the matrix

$$(A^{-1})^* = (A^*)^{-1}.$$

It is seen from (9) that x'_l is expressed in terms of $x_1, \dots, x_n$ with the aid of the l th column of matrix A^{-1} or the l th row of matrix $(A^{-1})^*$ conjugate to A^{-1} .

Further, it is seen from (6)

$$x_s = \sum_{i=1}^n a_{is} x'_i \quad (s = 1, \dots, n)$$

that the transition from $(x'_1, \dots, x'_n)$ to $(x_1, \dots, x_n)$ is accomplished with the aid of matrix A^* conjugate to A , i.e. x_s is expressed in terms of $x'_1, \dots, x'_n$ by means of the s th row of matrix A^* or the s th column of matrix A .

Remark. It was established in Sec. 14 that an arbitrary square matrix

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \quad (10)$$

determines the linear operator $y = Ax$ ($x \in R_n$, $y \in R_n$) defined by the formula

$$y_k = \sum_{j=1}^n a_{kj} x_j \quad (k = 1, \dots, n). \quad (11)$$

But the converse statement is also true: whatever the linear operator $y = Ax$ ($x \in R_n$, $y \in R_n$) may be, it is determined by a certain matrix (10) so that the vector $y = Ax$ is computed via the vector x by formulas (11).

Indeed, let there be given an arbitrary linear operator $y = Ax$ ($x \in R_n$, $y \in R_n$). Let us denote with its aid the images of the basis vectors i^s in the following way:

$$A(i^s) = (a_{1s}, a_{2s}, \dots, a_{ns}) = \sum_{k=1}^n a_{ks} i^k \quad (s = 1, \dots, n).$$

Then, by virtue of linearity of A , any vector

$$x = x_1 i^1 + \dots + x_n i^n = \sum_{s=1}^n x_s i^s$$

is mapped with the aid of A into the vector y defined by the equalities

$$\begin{aligned} y = Ax &= A\left(\sum_{s=1}^n x_s i^s\right) = \sum_{s=1}^n x_s A(i^s) \\ &= \sum_{s=1}^n x_s \sum_{k=1}^n a_{ks} i^k = \sum_{k=1}^n \left(\sum_{s=1}^n a_{ks} x_s\right) i^k, \end{aligned}$$

whence it follows that the k th component of $\mathbf{y}$ is determined by (11). Thus, operator A generates matrix (10) whose columns are formed by the coordinates of the basis vectors.

Example 1. Find the matrix of a linear operator (transformation) A consisting in rotation of the vectors of

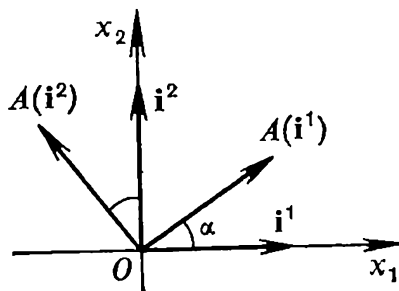


Fig. 33.

plane R_2 emanating from the origin through an angle α ($0 < \alpha < \pi/2$) anticlockwise.

Let us take the vectors $\mathbf{i}^1 = (1, 0)$, $\mathbf{i}^2 = (0, 1)$ for the basis. It is then obvious that (Fig. 33)

$$A(\mathbf{i}^1) = (\cos \alpha, \sin \alpha),$$

$$A(\mathbf{i}^2) = (-\sin \alpha, \cos \alpha).$$

Therefore, the matrix of our operator has the form

$$A = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}.$$

Sec. 17. Orthogonal Bases in R_n

Two nonzero vectors $\mathbf{x}, \mathbf{y} \in R_n$ are said to have *one and the same direction* if there exists a *positive* number λ such that $\mathbf{x} = \lambda \mathbf{y}$.

We say that an arbitrary nonzero vector $\mathbf{x} \in R_n$ can be *normalized* by replacing it by the unit vector

$$\mathbf{y} = \frac{1}{|\mathbf{x}|} \mathbf{x} \quad (|\mathbf{y}| = 1),$$

having the same direction as the vector $\mathbf{x}$.

The unit vector (having the norm (length) equal to 1) is called *normal*.

Two vectors $\mathbf{x}$ and $\mathbf{y}$ in space R_n are termed *orthogonal* if their scalar product is equal to zero: $(\mathbf{x}, \mathbf{y}) = 0$.

A system of vectors

$$\mathbf{x}^1, \dots, \mathbf{x}^v \in R_n \quad (1)$$

is called *orthogonal* if any two of its vectors are orthogonal. System of vectors (1) is said to be *orthogonal* and *normal* or, which is the same, *orthonormalized* if

$$(\mathbf{x}^k, \mathbf{x}^l) = \delta_{kl} = \begin{cases} 1, & k=l \\ 0, & k \neq l \end{cases} \quad (k, l = 1, \dots, v),$$

i.e. all the vectors of the system are normal and pairwise orthogonal. If system of vectors (1) is orthogonal and all of its vectors are not equal to zero, then, by normalizing them, we shall, obviously, get an orthonormalized system. *Orthonormalized system (1) is linearly independent*. Indeed, let

$$\lambda_1 \mathbf{x}^1 + \lambda_2 \mathbf{x}^2 + \dots + \lambda_v \mathbf{x}^v = \mathbf{0},$$

where $\lambda_1, \dots, \lambda_v$ are numbers. Multiplying this equation by $\mathbf{x}^s$ in a scalar way, we, obviously, get

$$\lambda_s (\mathbf{x}^s, \mathbf{x}^s) = \lambda_s = 0 \quad (s = 1, \dots, v).$$

But then an orthonormalized system of n vectors in R_n is a basis and, consequently, every vector $\mathbf{a} \in R_n$ can be represented in the form of a linear combination

$$\mathbf{a} = \sum_{k=1}^n \lambda_k \mathbf{x}^k. \quad (2)$$

Multiplying this equality by $\mathbf{x}^s$ in a scalar way, we get

$$(\mathbf{a}, \mathbf{x}^s) = \lambda_s \quad (s = 1, \dots, n)$$

and, consequently,

$$\mathbf{a} = \sum_{k=1}^n (\mathbf{a}, \mathbf{x}^k) \mathbf{x}^k, \quad \forall \mathbf{a} \in R_n.$$

We see that, conversely, orthogonality of matrix (8) implies orthonormalizability of the system of vectors $\mathbf{a}^1, \dots, \mathbf{a}^n$ defined by formulas (7).

This shows that *formulas (7), where $\|a_{ks}\|$ are arbitrary orthogonal matrices, determine all possible orthonormalized bases in R_n .*

Let us multiply equation (7) by $\mathbf{i}^l$ in a scalar manner:

$$(\mathbf{a}^k, \mathbf{i}^l) = a_{kl}.$$

Hence

$$\mathbf{i}^l = \sum_{k=1}^n (\mathbf{a}^k, \mathbf{i}^l) \mathbf{a}^k = \sum_{k=1}^n a_{kl} \mathbf{a}^k. \quad (11)$$

Thus, the transition from the basis $(\mathbf{a}^1, \dots, \mathbf{a}^n)$ to basis $(\mathbf{i}^1, \dots, \mathbf{i}^n)$ is realized with the aid of matrix Λ^* conjugate to Λ . Since transformations (11) are inverse to transformations (7) (see Sec. 15), we have simultaneously proved that the orthogonal matrix Λ possesses the following remarkable property:

$$\Lambda^{-1} = \Lambda^*.$$

It follows from (11):

$$\begin{aligned} \delta_{kl} = (\mathbf{i}^k, \mathbf{i}^l) &= \left(\sum_{s=1}^n a_{sk} \mathbf{a}^s, \sum_{r=1}^n a_{rl} \mathbf{a}^r \right) \\ &= \sum_{s=1}^n \sum_{r=1}^n a_{sk} a_{rl} (\mathbf{a}^s, \mathbf{a}^r) = \sum_{s=1}^n a_{sk} a_{sl}. \end{aligned} \quad (12)$$

We defined the orthogonal matrix as such a matrix in which the rows (i.e. the vectors representing them) are normal, and different rows are orthogonal. From this definition, as is seen from (12), it follows automatically that in an orthogonal matrix the columns are normal as well, and different columns are orthogonal.

The transition from $(x_1, \dots, x_n)$ to $(x'_1, \dots, x'_n)$ is effected with the aid of the matrix (see (9) in the preceding section, $\mathbf{x}' = \Lambda \mathbf{x}$)

$$(\Lambda^*)^{-1} = \Lambda^{**} = \Lambda,$$

i. e. (regarding that $\mathbf{a} = \sum_{l=1}^n x'_l \mathbf{a}^l$)

$$x'_l = \sum_{s=1}^n a_{ls} x_s. \quad (13)$$

And the transition from $(x'_1, \dots, x'_n)$ to $(x_1, \dots, x_n)$ is carried out via Λ^* conjugate to Λ (see (6) in the preceding section), i.e.

$$x_s = \sum_{l=1}^n a_{ls} x'_l.$$

Note that the determinant of the arbitrary orthogonal matrix Λ (see (6)) is equal (by absolute value) to 1: $\|\Lambda\| = \|a_{kl}\| = 1$.

This follows from the fact that

$$|\Lambda|^2 = |a_{kl}| \cdot |a_{ij}| = \left| \sum_{s=1}^n a_{ks} a_{is} \right| = \begin{vmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{vmatrix} = 1.$$

Here we hold that the element γ_{ki} of the product of the determinants is equal to the sum of products of the elements of the k th row by the corresponding elements of the i th row (see Sec. 2, property (j)).

Note also that the determinant formed of the components of the vectors constituting the basis $\mathbf{i}^1, \dots, \mathbf{i}^n$ is equal to 1:

$$\begin{vmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{vmatrix} = 1.$$

If an orthogonal basis $\mathbf{a}^1, \dots, \mathbf{a}^n$ has the determinant $|\Lambda| = 1$ (see (6)), then it is said to be *oriented in the same way as the basis $\mathbf{i}^1, \dots, \mathbf{i}^n$* . But if $|\Lambda| = -1$, then it is said to be *oriented in an opposite way*. These definitions conform to the corresponding definitions given in Sec. 11 for the two-dimensional and three-dimensional cases.

Remark. If in the matrix Λ (see (8)) the coordinates of the vector $\mathbf{a}^k$ in the basis $(\mathbf{i}^1, \dots, \mathbf{i}^n)$ were placed in the k th column ($k = 1, \dots, n$), then the transition from the coordinates of the vector $\mathbf{x}'$ to the coordinates of the vector $\mathbf{x}$ would be realized by means of the rows of matrix Λ . And the passage from $\mathbf{x}$ to $\mathbf{x}'$ in formula (13) would be achieved with the aid of matrix Λ^* .

Sec. 18. Invariant Properties of a Scalar and a Vector Product

Remark. In a three-dimensional space, let there be given two rectangular coordinate systems (x_1, x_2, x_3) and (y_1, y_2, y_3) with the systems of basis vectors $(\mathbf{i}^1, \mathbf{i}^2, \mathbf{i}^3)$ and $(\mathbf{j}^1, \mathbf{j}^2, \mathbf{j}^3)$, respectively. Let

$$\mathbf{j}^k = \sum_{s=1}^3 \alpha_{ks} \mathbf{i}^s \quad (k = 1, 2, 3) \quad (1)$$

(cf. (7) and (13), Sec. 17). Then the matrix

$$A = \|\alpha_{ks}\|$$

is orthogonal and (cf. (13), Sec. 17)

$$y_l = \sum_{s=1}^3 \alpha_{ls} x_s \quad (l = 1, 2, 3). \quad (2)$$

One and the same point (vector) of space has coordinates (components) (x_1, x_2, x_3) in the first coordinate system, and (y_1, y_2, y_3) in the second, the formulas for transforming the coordinates (i.e. for transition from (x_1, x_2, x_3) to (y_1, y_2, y_3)) being just the same as the formulas for transforming the system of basis vectors $(\mathbf{i}^1, \mathbf{i}^2, \mathbf{i}^3)$ to the system of basis vectors $(\mathbf{j}^1, \mathbf{j}^2, \mathbf{j}^3)$. In both cases we apply one and the same orthogonal matrix (see (13) and (7) in the preceding section)

$$A = \|\alpha_{kl}\|.$$

In the first case the matrix A is applied to the system of numbers (x_1, x_2, x_3) in order to obtain the system of numbers (y_1, y_2, y_3) , while in the second case the same matrix A is applied to the basis vectors $(\mathbf{i}^1, \mathbf{i}^2, \mathbf{i}^3)$ to obtain the basis vectors $(\mathbf{j}^1, \mathbf{j}^2, \mathbf{j}^3)$.

Let us give a general definition of the vector used in tensor calculus.

The vector in a three-dimensional Euclidean space is an entity that can be expressed in any rectangular coordinate system by a certain triple of numbers which are transformed in the same way (with the aid of the same matrix) as the triples of the basis vectors of the corresponding coordinate systems.

Similar terminology is also used for the case of n -dimensional spaces.

Let $\mathbf{a}$ and $\mathbf{b}$ be vectors of a three-dimensional space defined in coordinate systems with the basis vectors $(\mathbf{i}^1, \mathbf{i}^2, \mathbf{i}^3)$ and $(\mathbf{j}^1, \mathbf{j}^2, \mathbf{j}^3)$ as follows:

$$\mathbf{a} = a_{x_1}\mathbf{i}^1 + a_{x_2}\mathbf{i}^2 + a_{x_3}\mathbf{i}^3 = a_{y_1}\mathbf{j}^1 + a_{y_2}\mathbf{j}^2 + a_{y_3}\mathbf{j}^3,$$

$$\mathbf{b} = b_{x_1}\mathbf{i}^1 + b_{x_2}\mathbf{i}^2 + b_{x_3}\mathbf{i}^3 = b_{y_1}\mathbf{j}^1 + b_{y_2}\mathbf{j}^2 + b_{y_3}\mathbf{j}^3.$$

The following equality is valid:

$$\mathbf{ab} = \sum_{s=1}^3 a_{x_s} b_{x_s} = \sum_{r=1}^3 a_{y_r} b_{y_r},$$

showing that the *scalar product is invariant with respect to the transformations of rectangular coordinate systems.*

Indeed, since the systems of vectors $(\mathbf{i}^1, \mathbf{i}^2, \mathbf{i}^3)$ and $(\mathbf{j}^1, \mathbf{j}^2, \mathbf{j}^3)$ are orthonormalized, they are transformed by formulas (1), where $\|\alpha_{kl}\|$ is a certain orthogonal matrix. The components of the vector $\mathbf{a}$ ($a_{x_1}, a_{x_2}, a_{x_3}$) are transformed to the components ($a_{y_1}, a_{y_2}, a_{y_3}$) with the aid of the same matrix (see (2)). Therefore

$$\begin{aligned} \sum_{r=1}^3 a_{y_r} b_{y_r} &= \sum_{r=1}^3 \left(\sum_{s=1}^3 \alpha_{rs} a_{x_s} \cdot \sum_{v=1}^3 \alpha_{rv} b_{x_v} \right) \\ &= \sum_{s=1}^3 \sum_{v=1}^3 a_{x_s} b_{x_v} \left(\sum_{r=1}^3 \alpha_{rs} \alpha_{rv} \right) = \sum_{s=1}^3 a_{x_s} b_{x_s} = \mathbf{ab}. \end{aligned} \quad (3)$$

Thus, we have proved the invariance of the scalar product $\mathbf{ab}$ using a computational method. However, from the other, geometrical, definition of the scalar product of two vectors (directed line-segments), by virtue of which $\mathbf{ab} = |\mathbf{b}| \cdot \text{pr}_{\mathbf{b}} \mathbf{a}$, it is directly seen that this number is

an invariant, since this definition is not related with any coordinate system.

As regards the vector product $\mathbf{a} \times \mathbf{b}$, it is invariant relative to rectangular coordinate systems with common orientation. In systems having the basis vectors ($\mathbf{i}^1, \mathbf{i}^2, \mathbf{i}^3$) and ($\mathbf{j}^1, \mathbf{j}^2, \mathbf{j}^3$), the vector products are expressed in the following manner:

$$[\mathbf{a} \times \mathbf{b}]_{\mathbf{i}^1, \mathbf{i}^2, \mathbf{i}^3} = \begin{vmatrix} \mathbf{i}^1 & \mathbf{i}^2 & \mathbf{i}^3 \\ a_{x_1} & a_{x_2} & a_{x_3} \\ b_{x_1} & b_{x_2} & b_{x_3} \end{vmatrix},$$

$$[\mathbf{a} \times \mathbf{b}]_{\mathbf{j}^1, \mathbf{j}^2, \mathbf{j}^3} = \begin{vmatrix} \mathbf{j}^1 & \mathbf{j}^2 & \mathbf{j}^3 \\ a_{y_1} & a_{y_2} & a_{y_3} \\ b_{y_1} & b_{y_2} & b_{y_3} \end{vmatrix}.$$

By formulas (1) and (2),

$$[\mathbf{a} \times \mathbf{b}]_{\mathbf{j}^1, \mathbf{j}^2, \mathbf{j}^3} = \begin{vmatrix} \sum_{s=1}^3 \alpha_{1s} \mathbf{i}^s & \sum_{s=1}^3 \alpha_{2s} \mathbf{i}^s & \sum_{s=1}^3 \alpha_{3s} \mathbf{i}^s \\ \sum_{s=1}^3 \alpha_{1s} a_{x_s} & \sum_{s=1}^3 \alpha_{2s} a_{x_s} & \sum_{s=1}^3 \alpha_{3s} a_{x_s} \\ \sum_{s=1}^3 \alpha_{1s} b_{x_s} & \sum_{s=1}^3 \alpha_{2s} b_{x_s} & \sum_{s=1}^3 \alpha_{3s} b_{x_s} \end{vmatrix}$$

$$= \begin{vmatrix} \mathbf{i}^1 & \mathbf{i}^2 & \mathbf{i}^3 \\ a_{x_1} & a_{x_2} & a_{x_3} \\ b_{x_1} & b_{x_2} & b_{x_3} \end{vmatrix} \begin{vmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{vmatrix} = \pm \begin{vmatrix} \mathbf{i}^1 & \mathbf{i}^2 & \mathbf{i}^3 \\ a_{x_1} & a_{x_2} & a_{x_3} \\ b_{x_1} & b_{x_2} & b_{x_3} \end{vmatrix}$$

$$= \pm [\mathbf{a} \times \mathbf{b}]_{\mathbf{i}^1, \mathbf{i}^2, \mathbf{i}^3}, \quad (4)$$

where the sign $+$ or $-$ is taken depending on whether the determinant $|\alpha_{ik}|$ equals $+1$ or -1 , or, which is the same, on whether the transformation in question changes orientation.

In this case the following rule has been used for multiplying the determinants: the element γ_{ik} of the matrix of the product $\|\gamma_{ik}\|$ is determined as the product of the i th row of the first determinant by the k th row of the second (see Sec. 2, property (j)).

Thus, we have proved by a computational method that the *vector product of two vectors is invariant with respect to the transformations of rectangular coordinate systems which do not change their orientation.*

Transformations (3) and (4) are of interest as they are generalized for the case when the role of the vector $\mathbf{a}$ is played by the symbolic vector $\nabla = \left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \frac{\partial}{\partial x_3} \right)$ which is of great importance in mathematical analysis.

Sec. 19. Transformation of Rectangular Coordinates in a Plane

Let there be given a plane R_2 in which a rectangular coordinate system (x_1, x_2) is specified. Let then

$$\mathbf{i}^1 = (1, 0), \quad \mathbf{i}^2 = (0, 1)$$

be basis vectors of the axes x_1, x_2 . The basis vectors $\mathbf{i}^1, \mathbf{i}^2$ form an orthonormalized basis in R_2 .

An arbitrary unit (normal) vector $\mathbf{b}^1$ can be written in the following way:

$$\mathbf{b} = (\cos \alpha, \sin \alpha) \quad (0 \leq \alpha < 2\pi).$$

The unit vector orthogonal (perpendicular) to $\mathbf{b}^1$ will be denoted by $\mathbf{b}^2$. It may only correspond either to the angle $\alpha + \frac{\pi}{2}$, or $\alpha - \frac{\pi}{2}$. Since

$$\cos \left(\alpha + \frac{\pi}{2} \right) = -\sin \alpha, \quad \sin \left(\alpha + \frac{\pi}{2} \right) = \cos \alpha,$$

$$\cos \left(\alpha - \frac{\pi}{2} \right) = \sin \alpha, \quad \sin \left(\alpha - \frac{\pi}{2} \right) = -\cos \alpha,$$

all possible orthonormalized systems $\mathbf{b}^1, \mathbf{b}^2$ in R_2 are defined either by the equalities (Fig. 34)

$$\left. \begin{aligned} \mathbf{b}^1 &= \mathbf{i}^1 \cos \alpha + \mathbf{i}^2 \sin \alpha, \\ \mathbf{b}^2 &= -\mathbf{i}^1 \sin \alpha + \mathbf{i}^2 \cos \alpha \end{aligned} \right\} \quad (0 \leq \alpha < 2\pi), \quad (1')$$

corresponding to the rotation of the axes about the origin through an angle α with the orientation retained, or by

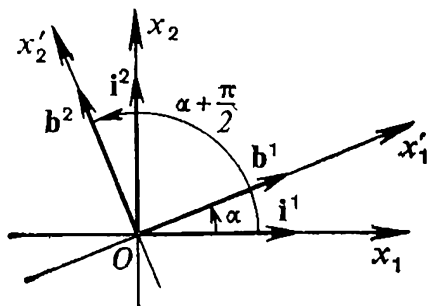


Fig. 34.

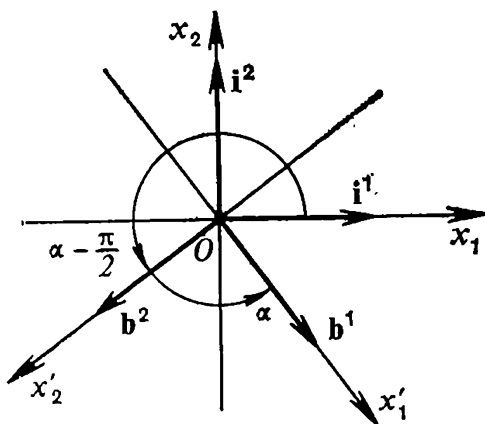


Fig. 35.

the equalities (Fig. 35)

$$\left. \begin{aligned} b^1 &= i^1 \cos \alpha + i^2 \sin \alpha, \\ b^2 &= i^1 \sin \alpha + i^2 \cos \alpha, \end{aligned} \right\} \quad (1'')$$

corresponding to the rotation of the axes about the origin through an angle α with the orientation changed.

Both transformations are united in the following formula:

$$\begin{aligned} b^1 &= \alpha_{11} i^1 + \alpha_{12} i^2, \\ b^2 &= \alpha_{21} i^1 + \alpha_{22} i^2, \end{aligned} \quad (1)$$

where the transformation matrix

$$\Lambda = \begin{vmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{vmatrix} \quad (2)$$

is orthogonal (the sum of the squared elements of any of its rows or columns is equal to 1, while the scalar product of two different rows or columns is zero).

Any defined orthogonal transformation (1) is, in fact, one of transformations (1'), (1'') for a certain α .

By virtue of the orthogonality of matrix (2), it follows from (1) that

$$\left. \begin{aligned} \mathbf{i}^1 &= \alpha_{11}\mathbf{b}^1 + \alpha_{21}\mathbf{b}^2, \\ \mathbf{i}^2 &= \alpha_{12}\mathbf{b}^1 + \alpha_{22}\mathbf{b}^2, \end{aligned} \right\} \quad (3)$$

and, thus, we have obtained a transformation inverse to transformation (1) with the matrix

$$\left\| \begin{array}{cc} \alpha_{11} & \alpha_{21} \\ \alpha_{12} & \alpha_{22} \end{array} \right\| = \Lambda^*,$$

which is conjugate to Λ .

Let there be given in a plane an arbitrary vector (point) $\mathbf{a}$ having the coordinates (x_1, x_2) and (x'_1, x'_2) in the old and new coordinate systems, respectively. Then

$$\mathbf{a} = x_1\mathbf{i}^1 + x_2\mathbf{i}^2 = x'_1\mathbf{b}^1 + x'_2\mathbf{b}^2. \quad (4)$$

By formulas (3) and (4)

$$\begin{aligned} x'_1\mathbf{b}^1 + x'_2\mathbf{b}^2 &= x_1(\alpha_{11}\mathbf{b}^1 + \alpha_{21}\mathbf{b}^2) + x_2(\alpha_{12}\mathbf{b}^1 + \alpha_{22}\mathbf{b}^2) \\ &= (\alpha_{11}x_1 + \alpha_{12}x_2)\mathbf{b}^1 + (\alpha_{21}x_1 + \alpha_{22}x_2)\mathbf{b}^2. \end{aligned}$$

Therefore, equating the components of identical basis vectors $\mathbf{b}^1, \mathbf{b}^2$, we get

$$\left. \begin{aligned} x'_1 &= \alpha_{11}x_1 + \alpha_{12}x_2, \\ x'_2 &= \alpha_{21}x_1 + \alpha_{22}x_2. \end{aligned} \right\} \quad (5)$$

And, by formulas (1) and (4),

$$\begin{aligned} x_1\mathbf{i}^1 + x_2\mathbf{i}^2 &= x'_1(\alpha_{11}\mathbf{i}^1 + \alpha_{12}\mathbf{i}^2) + x'_2(\alpha_{21}\mathbf{i}^1 + \alpha_{22}\mathbf{i}^2) \\ &= (\alpha_{11}x'_1 + \alpha_{21}x'_2)\mathbf{i}^1 + (\alpha_{12}x'_1 + \alpha_{22}x'_2)\mathbf{i}^2, \end{aligned}$$

whence, equating the components of $\mathbf{i}^1$ and $\mathbf{i}^2$, we get the formulas inverse to (5):

$$\left. \begin{aligned} x_1 &= \alpha_{11}x'_1 + \alpha_{21}x'_2, \\ x_2 &= \alpha_{12}x'_1 + \alpha_{22}x'_2. \end{aligned} \right\} \quad (6)$$

If, in addition to transformation (6), the origin of the axes x'_1, x'_2 is transferred to the point O' having the coordinates $x_1 = x_1^0, x_2 = x_2^0$, then formulas (6) will, obviously, be complicated as follows

$$\left. \begin{aligned} x_1 &= x_1^0 + \alpha_{11}x'_1 + \alpha_{21}x'_2, \\ x_2 &= x_2^0 + \alpha_{12}x'_1 + \alpha_{22}x'_2. \end{aligned} \right\} \quad (7)$$

Thus, an arbitrary transformation of the rectangular coordinates (x_1, x_2) to rectangular coordinates (x'_1, x'_2) , with the origin of the coordinate system (x_1, x_2) transferred to the point $O' = (x_1^0, x_2^0)$, is expressed by (7), where the matrix

$$\begin{vmatrix} \alpha_{11} & \alpha_{21} \\ \alpha_{12} & \alpha_{22} \end{vmatrix}$$

is orthogonal.

The corresponding transformation retaining the orientation of the given coordinate system has the form

$$\left. \begin{aligned} x_1 &= x_1^0 + x'_1 \cos \alpha - x'_2 \sin \alpha, \\ x_2 &= x_2^0 + x'_1 \sin \alpha + x'_2 \cos \alpha \end{aligned} \right\} \quad (7')$$

and the transformation, changing the orientation, has the form

$$\left. \begin{aligned} x_1 &= x_1^0 + x'_1 \cos \alpha + x'_2 \sin \alpha, \\ x_2 &= x_2^0 + x'_1 \sin \alpha - x'_2 \cos \alpha \end{aligned} \right\} \quad (7'')$$

(the matrices of the coefficients of x'_1 and x'_2 in (7') and (7'') transpose (1') and (1''), respectively).

Sec. 20. Linear Subspaces in R_n

The set L in R_n ($L \subset R_n$) is called a *linear subspace of the space R_n* , or, in brief, a subspace in R_n if from the fact that two certain vectors $\mathbf{x}$ and $\mathbf{y}$ belong to L ($\mathbf{x}, \mathbf{y} \in L$) it follows automatically that the vector $\alpha\mathbf{x} + \beta\mathbf{y}$ also belongs to L ($\alpha\mathbf{x} + \beta\mathbf{y} \in L$), where α, β are numbers. The subspace L is called *m-dimensional* if there is a linearly independent system $\mathbf{a}^1, \dots, \mathbf{a}^m$ in it consisting of m vectors, and there is no system consisting of $m + 1$ linearly independent vectors.

Here, if $\mathbf{a}$ is an arbitrary vector in L ($\mathbf{a} \in L$), then the system $\mathbf{a}^1, \dots, \mathbf{a}^m$, $\mathbf{a}$ is linearly dependent, i.e. there exists a nontrivial system of numbers $\lambda_1, \dots, \lambda_m, \lambda_{m+1}$ such that

$$\lambda_1 \mathbf{a}^1 + \dots + \lambda_m \mathbf{a}^m + \lambda_{m+1} \mathbf{a} = 0. \quad (1)$$

Here $\lambda_{m+1} \neq 0$ otherwise it would be

$$\lambda_1 \mathbf{a}^1 + \dots + \lambda_m \mathbf{a}^m = 0$$

and in consequence of the linear independence of the system $\mathbf{a}^1, \dots, \mathbf{a}^m$ it would be $\lambda_1 = \dots = \lambda_m = 0$, the entire system $\lambda_1, \dots, \lambda_m, \lambda_{m+1}$ would also be trivial. Then equation (1) can be solved with respect to $\mathbf{a}$:

$$\mathbf{a} = \mu_1 \mathbf{a}^1 + \dots + \mu_m \mathbf{a}^m \quad (\mu_s = -\lambda_s/\lambda_{m+1}), \quad (2)$$

i.e. can be represented in the form of a linear combination of the vectors $\mathbf{a}^1, \dots, \mathbf{a}^m$. On the other hand, the linear combination of form (2) belongs to L , since L is a subspace. In this sense, the system $\mathbf{a}^1, \dots, \mathbf{a}^m$ is said to be a *basis* in L . Obviously, any other linearly independent system of vectors $\mathbf{b}^1, \dots, \mathbf{b}^m$ belonging to L is a basis in L .

Decomposing the vectors $\mathbf{b}^i$ into the vectors $\mathbf{a}^1, \dots, \mathbf{a}^m$, we get

$$\mathbf{b}^k = \sum_{s=1}^m b_{ks} \mathbf{a}^s \quad (k=1, \dots, m).$$

Reasoning in the same way as in Sec. 16 for R_n (where now $\mathbf{i}^s$ and $\mathbf{a}^k$ have to be replaced by $\mathbf{a}^s$ and $\mathbf{b}^k$, respectively), we can obtain that the system $\mathbf{b}^1, \dots, \mathbf{b}^m$ is linearly independent when and only when the determinant $|b_{ks}| \neq 0$, and that any independent system consisting of $l < m$ vectors cannot be a basis in L .

The space R_n may be regarded as a subspace R_n having n dimensions.

The set consisting of one zero vector $\mathbf{0}$ is a linear subspace ($\alpha \mathbf{0} + \beta \mathbf{0} = \mathbf{0}$). It is said to have 0 dimensions. The vector $\mathbf{0}$ does not form a linearly independent system, since from the equality $\lambda \mathbf{0} = \mathbf{0}$, where λ is a number, does not necessarily follow that λ is zero.

If the vector $\mathbf{x}^0 \neq \mathbf{0}$, then the set of vectors of the form $\lambda \mathbf{x}^0$, where λ is an arbitrary number, is a *one-dimensional subspace*. The vector $\mathbf{x}^0$ may be taken as a basis in it.

Let L be a linear subspace in R_n . The vector $\mathbf{v} \in R_n$ will be said to be *orthogonal* to L if it is orthogonal to any vector $\mathbf{u} \in L$. Let us denote by L' the set of all vectors orthogonal to L . L' is a subspace. Indeed, let $\mathbf{v}, \mathbf{v}' \in L'$, i.e.

$$(\mathbf{v}, \mathbf{u}) = 0, \quad \forall \mathbf{u} \in L;$$

$$(\mathbf{v}', \mathbf{u}) = 0, \quad \forall \mathbf{u} \in L.$$

Then for any numbers α, β

$$(\alpha \mathbf{v} + \beta \mathbf{v}', \mathbf{u}) = \alpha (\mathbf{v}, \mathbf{u}) + \beta (\mathbf{v}', \mathbf{u}) = 0, \quad \forall \mathbf{u} \in L,$$

i.e. $\alpha \mathbf{v} + \beta \mathbf{v}' \in L'$.

The subspace $L' \subset R_n$, is termed *orthogonal to the given subspace $L \subset R_n$* if L' is the set of all vectors each of which is orthogonal to L .

Proved below is a theorem explaining the structure of an arbitrary subspace $L \subset R_n$ and the subspace $L' \subset R_n$ orthogonal to it. In particular, it follows from this theorem that if L' is orthogonal to L , then the converse is also true: L is orthogonal to L' .

Theorem 1. *Let L be a linear subspace different from R_n and from a zero subspace. Then:*

(a) *there exists a whole number m satisfying the inequalities*

$$1 \leq m < n, \quad (3)$$

and an orthonormalized basis

$$\mathbf{a}^1, \dots, \mathbf{a}^m \quad (4)$$

in L ; if this basis is extended in any way to the basis in R_n

$$\mathbf{a}^1, \dots, \mathbf{a}^m, \mathbf{a}^{m+1}, \dots, \mathbf{a}^n, \quad (5)$$

then the linear subspace L' with the basis

$$\mathbf{a}^{m+1}, \dots, \mathbf{a}^n \quad (6)$$

possesses the following properties:

(b) L' is a subspace orthogonal to L ;

(c) L is a subspace orthogonal to L' ;

(d) any vector $\mathbf{a} \in R_n$ can be represented in the form of a sum

$$\mathbf{a} = \mathbf{u} + \mathbf{v},$$

where $\mathbf{u} \in L$, $\mathbf{v} \in L'$ and this can be done in a unique way.

Proof. By the hypothesis, L is different from a zero subspace, consequently, in L there exists a vector $\mathbf{x}$ different from $\mathbf{0}$. Normalizing $\mathbf{x}$, we get a normal vector

$$\mathbf{a}^1 = \frac{1}{|\mathbf{x}|} \mathbf{x}.$$

Let us denote by $\mathbf{a}^2$ any (belonging to L) normal vector orthogonal to $\mathbf{a}^1$ ($|\mathbf{a}^2| = 1$, $(\mathbf{a}^2, \mathbf{a}^1) = 0$), provided that such a vector exists. Further, let us denote by $\mathbf{a}^3$ a normal vector, belonging to L , that is orthogonal to $\mathbf{a}^1$ and $\mathbf{a}^2$ ($|\mathbf{a}^3| = 1$, $(\mathbf{a}^3, \mathbf{a}^1) = (\mathbf{a}^3, \mathbf{a}^2) = 0$), provided that such a vector exists. This process will be completed at a certain m th step, where m satisfies inequalities (3), i.e. we can find an orthonormalized system of vectors (4) belonging to L , but the latter will not contain a unit vector orthogonal to the vectors $\mathbf{a}^1, \dots, \mathbf{a}^m$. Indeed, $m \geq 1$, since, automatically, $\mathbf{a}^1 \in L$. On the other hand, m cannot be equal to n . Otherwise, the vectors $\mathbf{a}^1, \dots, \mathbf{a}^n$ would belong to L and, together with them, all linear combinations $\sum_{k=1}^n \lambda_k \mathbf{a}^k$ would belong to the subspace L .

And this would result in that L coincides with R_n , but L is different from R_n . The obtained orthonormalized system $\mathbf{a}^1, \dots, \mathbf{a}^m$ is a basis in L . Indeed, together with the vectors $\mathbf{a}^1, \dots, \mathbf{a}^m$, all their linear combinations $\sum_{k=1}^m \lambda_k \mathbf{a}^k$ belong to L . But there are no other vectors in L , since if we assume that a certain vector $\mathbf{a} \in L$ is not such a linear combination, then $\mathbf{a}$ could be written in the form of a sum

$$\mathbf{a} = \sum_{k=1}^m (\mathbf{a}, \mathbf{a}^k) \mathbf{a}^k + \mathbf{y}, \quad (7)$$

where $\mathbf{y} \neq 0$. Since the vectors $\mathbf{a}$ and $\mathbf{a}^k$ belong to the subspace L , we had to conclude that the vector

$$\mathbf{y} = \mathbf{a} - \sum_{k=1}^m (\mathbf{a}, \mathbf{a}^k) \mathbf{a}^k$$

also belongs to L . But the vector $\mathbf{y}$ is orthogonal to all $\mathbf{a}^s$ ($s = 1, \dots, m$) (see (4) in Sec. 17). The normalized vector

$$\mathbf{b} = \frac{1}{|\mathbf{y}|} \mathbf{y} \quad (8)$$

would also belong to L and would be orthogonal to all $\mathbf{a}^k$ ($k = 1, \dots, m$). But this is impossible due to the maximal property of the number m . This proves statement (a) of the theorem under consideration.

Orthonormalized system (4) is supplemented to orthonormalized basis (5) on the basis of Theorem 1 from Sec. 17. We denote by L' the subspace of all linear combinations

$\mathbf{v} = \sum_{k=m+1}^n \mu_k \mathbf{a}^k$ from the vectors of system (6). Every such vector is, obviously, orthogonal to any vector $\mathbf{u} \in L$

that is represented in the form of a sum $\mathbf{u} = \sum_{k=1}^m \lambda_k \mathbf{a}^k$.

On the other hand, if $\mathbf{a} \in R_n$ is an arbitrary vector orthogonal to all vectors $\mathbf{u} \in L$, to $\mathbf{a}^1, \dots, \mathbf{a}^m$ in particular, then its decomposition according to basis (5) has the form

$$\mathbf{a} = \sum_{k=1}^n (\mathbf{a}, \mathbf{a}^k) \mathbf{a}^k = \sum_{k=m+1}^n (\mathbf{a}, \mathbf{a}^k) \mathbf{a}^k,$$

i.e. $\mathbf{a} \in L'$. We have proved statement (b) of our theorem.

Further, any vector $\mathbf{u} = \sum_{k=1}^m \lambda_k \mathbf{a}^k \in L$ is orthogonal to all vectors $\mathbf{v} = \sum_{k=m+1}^n \mu_k \mathbf{a}^k \in L'$ and, if it is known that a certain vector $\mathbf{a} = \sum_{k=1}^n (\mathbf{a}, \mathbf{a}^k) \mathbf{a}^k$ is orthogonal to all vectors from L' , to $\mathbf{a}^{m+1}, \dots, \mathbf{a}^n$ in particular, then

$\mathbf{a} = \sum_{k=1}^m (\mathbf{a}, \mathbf{a}^k) \mathbf{a}^k$, that is $\mathbf{a} \in L$. And so, we have proved statement (c).

Finally, if $\mathbf{a} \in R_n$ is an arbitrary vector, then it can be represented in the form of a sum in a unique manner:

$$\mathbf{a} = \sum_{k=1}^n (\mathbf{a}, \mathbf{a}^k) \mathbf{a}^k = \mathbf{u} + \mathbf{v},$$

where

$$\mathbf{u} = \sum_{k=1}^m (\mathbf{a}, \mathbf{a}^k) \mathbf{a}^k \in L, \quad \mathbf{v} = \sum_{k=m+1}^n (\mathbf{a}, \mathbf{a}^k) \mathbf{a}^k \in L'.$$

Thus, Theorem 1 has been completely proved.

Theorem 2. *Let L be a subspace in R_n having m dimensions. Then the subspace $L' \subset R_n$ orthogonal to L has $n - m$ dimensions, and L is, in its turn, a subspace orthogonal to L' .*

Proof. If L is different both from R_n and zero subspace, then this theorem is, obviously, contained in Theorem 1. Rejecting from the formulation of Theorem 1 everything concerning bases (4), (5) and (6), we obtain Theorem 2.

Let L be a zero subspace. Since any vector $\mathbf{a} \in R_n$ is orthogonal to $\mathbf{0}$, $L' = R_n$ and the dimension of R_n is equal to $n - 0 = n$. Conversely, the vector $\mathbf{0}$ is orthogonal to all vectors, $\mathbf{a} \in R_n = L'$. There are no other vectors orthogonal to all vectors R_n , since any vector different from $\mathbf{0}$ is no longer orthogonal to itself. Thus, we have proved that L is orthogonal to L' .

If $L = R_n$, we resort to similar arguments.

Corollary 1. *Let there be given a system of vectors*

$$\mathbf{x}^1, \dots, \mathbf{x}^m, \tag{9}$$

and let L' be a subspace of vectors $\mathbf{v}$ each of which is orthogonal to the vectors of this system:

$$(\mathbf{v}, \mathbf{x}^k) = 0 \quad (k = 1, \dots, m).$$

Further, let there be given a vector $\mathbf{a}$ orthogonal to all indicated vectors $\mathbf{v}$, i.e. orthogonal to the subspace L' . Then $\mathbf{a}$ is a certain linear combination of the vectors of the given

system (9)

$$\mathbf{a} = \sum_{k=1}^m \lambda_k \mathbf{x}^k$$

Proof. Consider the subspace L consisting of all possible linear combinations of the vectors of system (9), i.e. any vector $\mathbf{u} \in L$ is a certain linear combination

$$\mathbf{u} = \sum_{k=1}^m \lambda_k \mathbf{x}^k.$$

In this case we shall also say that the subspace L is stretched on the vectors of system (8).

Since any vector $\mathbf{v} \in L'$ is orthogonal to the vectors of system (9), it is, obviously, orthogonal to any vector $\mathbf{u} \in L$. This shows that subspace L' is orthogonal to subspace L . But then, by Theorem 2, L is also orthogonal to L' , i.e. L consists of *all* vectors $\mathbf{u}$ orthogonal to L' . By hypothesis, $\mathbf{a}$ is one of such vectors $\mathbf{u}$, consequently, $\mathbf{a}$ is a certain linear combination of the vectors of system (9).

Sec. 21. Fredholm-type Theorems

This section deals with the theory of linear equations which is parallel to the theory discussed in Sec. 4.

This is a determinant-free theory. The determinant of a system of equations does not enter into its formulation explicitly. Its advantage consists in that it served as the foundation and analogy for many generalizations in mathematical analysis. The first important generalizations of this kind belong to E. Fredholm*.

We consider the linear operator A once again (see Sec. 15):

$$\mathbf{y} = A\mathbf{x} \quad (\mathbf{x} \in R_n). \quad (1)$$

It associates every vector $\mathbf{x} \in R_n$ with a vector $\mathbf{y} \in R_n$ with the aid of the equations

$$y_i = \sum_{s=1}^n a_{is} x_s \quad (i = 1, \dots, n). \quad (2)$$

* Fredholm, Erik Ivar (1866-1927). Swedish analyst and physicist.

Here

$$A = \| a_{is} \| = \left\| \begin{array}{cccc} a_{11} & a_{12} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{array} \right\| \quad (3)$$

is a given quadratic matrix. Operator A is associated with the conjugate operator

$$\mathbf{y} = A^* \mathbf{x} \quad (\mathbf{x} \in R_n), \quad (1^*)$$

defined by the matrix

$$A^* = \left\| \begin{array}{cccc} a_{11} & a_{21} & \dots & a_{n1} \\ \dots & \dots & \dots & \dots \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{array} \right\|, \quad (3^*)$$

which is conjugate to (2). With the aid of the components of the vectors $\mathbf{x}$, $\mathbf{y}$, it is written in the form:

$$y_j = \sum_{l=1}^n a_{lj} x_l \quad (j = 1, \dots, n), \quad (2^*)$$

i.e. the component y_j is expressed in terms of the coordinates of the vector $\mathbf{x}$ with the aid of the j th row of matrix A^* or the j th column of matrix A .

The following equality is valid

$$(A\mathbf{x}, \mathbf{z}) = (\mathbf{x}, A^*\mathbf{z}), \quad \forall \mathbf{x}, \mathbf{z} \in R_n, \quad (4)$$

which is true for all $\mathbf{x}, \mathbf{z} \in R_n$. Indeed,

$$\begin{aligned} (A\mathbf{x}, \mathbf{z}) &= \sum_{i=1}^n y_i z_i = \sum_{i=1}^n \left(\sum_{s=1}^n a_{is} x_s \right) z_i \\ &= \sum_{s=1}^n \left(\sum_{i=1}^n a_{is} z_i \right) x_s = (\mathbf{x}, A^*\mathbf{z}). \end{aligned}$$

Equation (4) is characteristic for a conjugate operator, since, if for some linear operator B the equality

$$(A\mathbf{x}, \mathbf{z}) = (\mathbf{x}, B\mathbf{z}), \quad \forall \mathbf{x}, \mathbf{z} \in R_n, \quad (5)$$

is fulfilled, then it is necessary that $B = A^*$. Indeed ($B = \|b_{ik}\|$).

$$(Ax, z) = \sum_{i=1}^n \sum_{s=1}^n a_{is} x_s z_i,$$

$$(x, Bz) = \sum_{s=1}^n \sum_{i=1}^n b_{si} z_i x_s = \sum_{i=1}^n \sum_{s=1}^n b_{si} x_s z_i.$$

It follows from (5) that

$$\sum_{i=1}^n \sum_{s=1}^n a_{is} x_s z_i = \sum_{i=1}^n \sum_{s=1}^n b_{si} x_s z_i, \quad \forall x, z \in R_n, \quad (6)$$

whence $a_{is} = b_{si}$ ($i, s = 1, \dots, n$), which can be verified if we set in (6): $x = (0, \dots, 0, 1, 0, \dots, 0)$ and $z = (0, \dots, 0, 1, 0, \dots, 0)$, where 1 occupies the s th place in x and the i th place in z .

Hence, operator A^* conjugate to the linear operator A can also be defined as such a linear operator for which equality (4) is fulfilled.

Equalities (1) and (1*) may be considered as equations—given a vector $y \in R_n$, and we find $x \in R_n$, for which either equality (1), or (1*) is fulfilled.

The corresponding homogeneous equations have the form

$$Ax = 0 \quad (1_0)$$

or

$$\sum_{s=1}^n a_{is} x_s = 0 \quad (i = 1, \dots, n) \quad (2_0)$$

and

$$A^*z = 0 \quad (1^*_0)$$

or

$$\sum_{l=1}^n a_{lj} z_l = 0 \quad (j = 1, \dots, n). \quad (2^*_0)$$

Let us denote by L the image of the space R_n with the aid of operator A :

$$L = A(R_n),$$

and by L' the subspace of all vectors z satisfying homogeneous conjugate equation (1*₀).

We have called L' a subspace, since, together with $\mathbf{z}$, $\mathbf{z}'$, it also contains $\alpha\mathbf{z} + \beta\mathbf{z}'$, where α and β are numbers:

$$A^*(\alpha\mathbf{z} + \beta\mathbf{z}') = \alpha A^*\mathbf{z} + \beta A^*\mathbf{z}' = \mathbf{0}.$$

L is also a subspace, since, if $\mathbf{y}$, $\mathbf{y}' \in L$, then there exist vectors $\mathbf{x}$, $\mathbf{x}' \in R_n$ such that $\mathbf{y} = A\mathbf{x}$, $\mathbf{y}' = A\mathbf{x}'$, and, consequently,

$$\alpha\mathbf{y} + \beta\mathbf{y}' = \alpha A\mathbf{x} + \beta A\mathbf{x}' = A(\alpha\mathbf{x} + \beta\mathbf{x}'),$$

i.e. $\alpha\mathbf{y} + \beta\mathbf{y}' \in L$.

The following lemma is valid (see Sec. 20, Theorem 2).

Lemma 1. *Subspaces L and L' are mutually orthogonal, i.e. L' is the set of all vectors $\mathbf{z}$ each of which is orthogonal to L , and L , in its turn, is the set of all vectors $\mathbf{y}$ each of which is orthogonal to L' . If L has k dimensions, then L' has $n - k$ dimensions.*

Proof. Consider the equality

$$(A\mathbf{x}, \mathbf{z}) = (\mathbf{x}, A^*\mathbf{z}), \quad (7)$$

which is true for all $\mathbf{x}$, $\mathbf{z} \in R_n$. Let $\mathbf{z}$ be a vector orthogonal to L , then, for it, the left-hand member of (7) is equal to zero for all $\mathbf{x} \in R_n$, but then the right-hand member is also equal to zero for all $\mathbf{x} \in R_n$, for $\mathbf{x} = A^*\mathbf{z}$, in particular:

$$(A^*\mathbf{z}, A^*\mathbf{z}) = 0.$$

Consequently, $A^*\mathbf{z} = \mathbf{0}$. We have proved that if the vector $\mathbf{z}$ is orthogonal to L , then it satisfies the equation $A^*\mathbf{z} = \mathbf{0}$ (i.e. $\mathbf{z} \in L'$).

Conversely, let the vector $\mathbf{z}$ satisfy the equation $A^*\mathbf{z} = \mathbf{0}$. For such a $\mathbf{z}$ the right-hand member of (7) equals zero for any $\mathbf{x}$, but then the left-hand member is also equal to zero, i.e. $\mathbf{z}$ is orthogonal to all vectors of the form $A\mathbf{x}$, i.e. to all vectors $\mathbf{y} \in L$. In other words, $\mathbf{z}$ is orthogonal to L .

Thus, we have proved that L' is the set of *all* vectors $\mathbf{z}$ orthogonal to subspace L . But then, by Theorem 2 of Sec. 20, and conversely, L is the set of *all* vectors $\mathbf{y}$ orthogonal to L' , and the sum of the dimensions of L and L' is equal to n . The lemma has been proved.

The following theorem holds true.

Theorem 1. *In order for the equation*

$$\mathbf{y} = A\mathbf{x} \quad (1')$$

to have a solution for a given vector $\mathbf{y} \in R_n$, it is necessary and sufficient that the vector $\mathbf{y}$ be orthogonal to all vectors $\mathbf{z}$ satisfying the homogeneous conjugate equation

$$A^*\mathbf{z} = 0 \quad (1'')$$

The solution $\mathbf{x}$ of equation (1), provided that it exists, can be written in the form of a sum

$$\mathbf{x} = \mathbf{x}^0 + \mathbf{u},$$

where $\mathbf{x}^0$ is a certain particular solution of equation (1), and $\mathbf{u}$ is an arbitrary solution of the homogeneous equation

$$A\mathbf{u} = 0 \quad (1''')$$

Any indicated sum is a solution of (1').

Proof. By virtue of Lemma 1, if $L = A(R_n)$ and L' is the set of all $\mathbf{z}$ satisfying the equation $A^*\mathbf{z} = 0$, then L and L' are mutually orthogonal subspaces. But then, if for $\mathbf{y}$ there exists a solution of equation (1), $\mathbf{y} \in L$ and it is necessary that all $\mathbf{z} \in L'$ be orthogonal to $\mathbf{y}$. But if the vector $\mathbf{y}$ is orthogonal to all $\mathbf{z} \in L'$, then $\mathbf{y} \in L$, i.e. there exists an $\mathbf{x}$ for which $\mathbf{y} = A\mathbf{x}$.

Let now there exist a solution of equation (1') for vector $\mathbf{y}$. We denote it by $\mathbf{x}^0$:

$$\mathbf{y} = A\mathbf{x}^0.$$

Then, obviously, the sum $\mathbf{x}^0 + \mathbf{u}$, where $A\mathbf{u} = 0$ is also a solution of equation (1'):

$$A(\mathbf{x}^0 + \mathbf{u}) = A\mathbf{x}^0 + A\mathbf{u} = \mathbf{y} + 0 = \mathbf{y}.$$

Conversely, if $\mathbf{x}$ is an arbitrary solution of equation (1'), and $\mathbf{x}^0$ is a definite particular solution, then

$$\mathbf{y} = A\mathbf{x}, \quad \mathbf{y} = A\mathbf{x}^0,$$

and, consequently,

$$0 = A\mathbf{x} - A\mathbf{x}^0 = A(\mathbf{x} - \mathbf{x}^0) = A\mathbf{u},$$

where $\mathbf{u} = \mathbf{x} - \mathbf{x}^0$, i.e. $\mathbf{x} = \mathbf{x}^0 + \mathbf{u}$, where $\mathbf{u}$ satisfies the equation $A\mathbf{u} = 0$.

Remark. Taking the space R_2 as an example, let us clarify the connection of Theorem 1 with the Kronecker-Capelli theory. Let the vector $\mathbf{y} = (y_1, y_2)$ be orthogonal to all the solutions of the system

$$\left. \begin{aligned} a_{11}z_1 + a_{21}z_2 &= 0, \\ a_{12}z_1 + a_{22}z_2 &= 0. \end{aligned} \right\} \quad (8)$$

Let us show that the ranks of matrix A and the augmented matrix

$$B = \left\| \begin{array}{ccc} a_{11} & a_{12} & y_1 \\ a_{21} & a_{22} & y_2 \end{array} \right\|$$

are then equal to each other. If $\text{rank } A = 2$, then, obviously, $\text{rank } B = 2$. Let $\text{rank } A = 1$. It is always true that $\text{rank } B \geq \text{rank } A = 1$. Therefore, we have to prove that

$$\Delta_1 = \begin{vmatrix} a_{11} & y_1 \\ a_{21} & y_2 \end{vmatrix} = 0, \quad \Delta_2 = \begin{vmatrix} a_{12} & y_1 \\ a_{22} & y_2 \end{vmatrix} = 0.$$

Indeed, since $\mathbf{y}$ is orthogonal to nontrivial solutions of system (8), we have $y_1z_1 + y_2z_2 = 0$. Therefore, holding that $z_1 \neq 0$,

$$\begin{aligned} \Delta_1 &= a_{11}y_2 - a_{21}y_1 \\ &= a_{11}y_2 + a_{21} \frac{y_2z_2}{z_1} = \frac{y_2}{z_1} (a_{11}z_1 + a_{21}z_2) = 0, \end{aligned}$$

$$\begin{aligned} \Delta_2 &= a_{12}y_2 - a_{22}y_1 \\ &= a_{12}y_2 + a_{22} \frac{y_2z_2}{z_1} = \frac{y_2}{z_1} (a_{12}z_1 + a_{22}z_2) = 0. \end{aligned}$$

Hence it follows that $\text{rank } B = \text{rank } A = 1$.

Conversely, let the vector $\mathbf{y} = (y_1, y_2)$ be such that $\text{rank } B = \text{rank } A$, then (1) has a certain solution (x_1, x_2) . We are going to prove that $\mathbf{y}$ is orthogonal to the solutions $\mathbf{z} = (z_1, z_2)$ of system (8). Indeed,

$$\begin{aligned} y_1z_1 + y_2z_2 &= (a_{11}x_1 + a_{12}x_2)z_1 + (a_{21}x_1 + a_{22}x_2)z_2 \\ &= (a_{11}z_1 + a_{21}z_2)x_1 + (a_{12}z_1 + a_{22}z_2)x_2 = 0 \cdot x_1 + 0 \cdot x_2 \\ &= 0. \end{aligned}$$

Theorem 2. *The homogeneous equations*

$$A\mathbf{x} = \mathbf{0} \quad (10)$$

and

$$A^*_{\mathbf{X}} = 0 \quad (1^*)$$

have an equal number of linearly independent solutions.

In particular, if one of these equations has only a trivial solution 0, i.e. a zero of independent solutions, then it is true for the other equation as well.

Remark. In the latter case equation (1) has a unique solution.

Proof. The matrices A and A^* have one and the same rank which will be denoted by k . They also have one and the same determinant Δ .

If $k = n$, then $\Delta \neq 0$ and equations (1₀) and (1₀^{*}) have only trivial solutions 0. In this case, according to Theorem 1, equation (1) has a unique solution for any $y \in R_n$.

Let now $1 \leq k < n$. On renumbering the equations and components, the determinant

$$\begin{vmatrix} a_{11} & \dots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{k1} & \dots & a_{kk} \end{vmatrix} \neq 0.$$

Let us now write the first k equations of (1⁰) in the form

[illegible]

Given below is a table of $n - k$ vectors:

$$\left. \begin{aligned} \mathbf{x}^1 &= (x_1^1, \dots, x_h^1, 1, 0, \dots, 0), \\ \mathbf{x}^2 &= (x_1^2, \dots, x_k^2, 0, 1, 0, \dots, 0), \\ &\vdots \\ \mathbf{x}^{n-h} &= (x_1^{n-h}, \dots, x_k^{n-h}, 0, \dots, 0, 1). \end{aligned} \right\} \quad (10)$$

In order to obtain the first vector, we substitute into system (9)

$$x_{k+1} = 1, \quad x_{k+2} = 0, \quad \dots, \quad x_n = 0$$

and solve it with respect to $x_1, \dots, x_k$. We shall denote the unique solutions obtained here by $x_1^1, \dots, x_k^1$. To get the second vector, we substitute into (9)

$$x_{k+1} = 0, \quad x_{k+2} = 1, \quad x_{k+3} = 0, \quad \dots, \quad x_n = 0$$

and find the numbers $x_1^2, \dots, x_k^2$, and so on. Vectors (10) possess the following properties:

(1) System of vectors (10) is linearly independent, since the rank of the matrix of these vectors is equal to the number of these vectors $\mu = n - k$.

(2) Every vector of system (10) is a solution of (any!) equations (1₀) or $Ax = 0$.

(3) All possible solutions of the equation $Ax = 0$ have the form

$$\lambda_1 x^1 + \dots + \lambda_{n-k} x^{n-k},$$

where $\lambda_1, \dots, \lambda_{n-k}$ are arbitrary numbers.

These three assertions are usually replaced by the words: equation (1₀) has $n - k$ linearly independent solutions.

Reasoning in such a way and taking into consideration that $\text{rank } A = \text{rank } A^*$, we prove that the equation $A^*x = 0$ also has $n - k$ linearly independent solutions. The theorem has been proved.

Theorem 3. *If one of homogeneous equations (1₀) or (1₀^{*}) has k linearly independent solutions, then the other also has k linearly independent solutions; and the images $L = A(R_n)$ and $L^* = A^*(R_n)$ of space R_n obtained with the aid of operators A and A^* are subspaces of $n - k$ dimensions.*

Proof. The first statement of this theorem concerning the equality of the numbers of linearly independent solutions of homogeneous equations (1₀) and (1₀^{*}) is Theorem 2, while the second is Lemma 1, by virtue of which the dimension of subspace L is equal to $n - k$, where k is the dimension of the subspace L' of the vectors z satisfying the equation $A^*z = 0$. Analogously, the dimension of L^* is equal to $n - k$, where k is the number of dimensions of the subspace of the vectors u satisfying the equation $Au = 0$.

Sec. 22. Self-adjoint Operator. Quadratic Form

A linear operator

$$y_k = \sum_{l=1}^n a_{kl} x_l \quad (k = 1, \dots, n), \quad (1)$$

where we considered a_{kl} to be real numbers, or, briefly,

$$y = Ax \quad (x \in R_n, \quad y \in R_n) \quad (2)$$

is called *self-adjoint* if it is equal to its conjugate operator ($A = A^*$), i.e. if

$$Ax = A^*x, \quad \forall x \in R_n, \quad (3)$$

or in other words, if matrix A is symmetric:

$$a_{kl} = a_{lk} \quad (k, l = 1, \dots, n) \quad (4)$$

(see (3) and (3*) in the preceding section).

For a self-adjoint operator the following characteristic equality holds true:

$$(x, Az) = (Ax, z), \quad \forall x, z \in R_n$$

(see (4) in the preceding section). Obviously,

$$(x, Ax) = \sum_{k=1}^n x_k \sum_{l=1}^n a_{kl} x_l = \sum_{k=1}^n \sum_{l=1}^n a_{kl} x_k x_l \quad (a_{kl} = a_{lk}). \quad (4')$$

The expression on the right in (4') is called the *quadratic form of the n th order*. This is a continuous function of the vector x or, which is the same, of the variables $x_1, \dots, x_n$.

We shall consider this function on the set S of values of x having the unit norm ($|x| = 1$). The set S is a sphere in R_n of radius 1 with centre at the point 0. S is a bounded set; besides, it is closed*: if the points of the sequence $\{x^v\}$ ($v = 1, 2, \dots$) belong to S (i.e. $|x^v| = 1$, $v = 1, 2, \dots$) and this sequence tends to a certain point $x^0 \in R_n$ ($x^v \rightarrow x^0$, $v \rightarrow \infty$), then, inevitably, $x^0 \in S$, i.e. $|x^0| = 1$, since $|1 - |x^0|| = ||x^v| - |x^0|| \leq |x^v - x^0| \rightarrow 0$, whence $|x^0| = 1$.

Let us find the maximum of quadratic form (4') on the sphere S . Since form (4') is a continuous function on a closed bounded set, its maximum on S is achieved for a certain unit vector x^1 ($|x^1| = 1$). Let us denote this

* See our book "Differential and Integral Calculus" (the second part of our series "Higher Mathematics"), Mir Publishers, Moscow, 1982, Sec. 8.12.

maximum by λ_1 :

$$\lambda_1 = (A\mathbf{x}^1, \mathbf{x}^1) \geq (A\mathbf{x}, \mathbf{x}), \quad \forall \mathbf{x} : |\mathbf{x}| = 1. \quad (5)$$

We then introduce a subspace L' orthogonal to the vector $\mathbf{x}^1$, i.e. the set of all vectors $\mathbf{v}$ each of which is orthogonal to $\mathbf{x}^1$, and take an arbitrary unit vector $\mathbf{v}^0$ ($|\mathbf{v}^0| = 1$) in L' . The vector

$$\cos \alpha \cdot \mathbf{x}^1 + \sin \alpha \cdot \mathbf{v}^0$$

depends on α and has a unit norm:

$$\begin{aligned} & |\cos \alpha \cdot \mathbf{x}^1 + \sin \alpha \cdot \mathbf{v}^0| \\ &= (\cos \alpha \cdot \mathbf{x}^1 + \sin \alpha \cdot \mathbf{v}^0, \cos \alpha \cdot \mathbf{x}^1 + \sin \alpha \cdot \mathbf{v}^0)^{1/2} \\ &= (\cos^2 \alpha + \sin^2 \alpha)^{1/2} = 1 \end{aligned}$$

For $\alpha = 0$ this vector turns into $\mathbf{x}^1$. But then the function

$$\psi(\alpha) = (A(\cos \alpha \cdot \mathbf{x}^1 + \sin \alpha \cdot \mathbf{v}^0), \cos \alpha \cdot \mathbf{x}^1 + \sin \alpha \cdot \mathbf{v}^0)$$

reaches its maximum at the point $\alpha = 0$ ($\psi(0) = (A\mathbf{x}^1, \mathbf{x}^1)$) and, by virtue of a necessary condition for an extremum,

$$\psi'(0) = 0.$$

Let us compute this derivative. We have

$$\begin{aligned} \psi(\alpha) &= \cos^2 \alpha (A\mathbf{x}^1, \mathbf{x}^1) + \sin 2\alpha (A\mathbf{x}^1, \mathbf{v}^0) \\ &\quad + \sin^2 \alpha (A\mathbf{v}^0, \mathbf{v}^0). \end{aligned}$$

Consequently,

$$\begin{aligned} \psi'(\alpha) &= -\sin 2\alpha (A\mathbf{x}^1, \mathbf{x}^1) + 2 \cos 2\alpha (A\mathbf{x}^1, \mathbf{v}^0) \\ &\quad + \sin 2\alpha (A\mathbf{v}^0, \mathbf{v}^0) \end{aligned}$$

and

$$\psi'(0) = 2 (A\mathbf{x}^1, \mathbf{v}^0) = 0.$$

We have obtained that the vector $A\mathbf{x}^1$ is orthogonal to all unit vectors $\mathbf{v}^0 \in L'$ and, consequently, to any vectors $\mathbf{v} \in L'$. But then $A\mathbf{x}^1$ differs from $\mathbf{x}^1$ only by a factor (see Corollary 1 at the end of Sec. 20), i.e.

$$A\mathbf{x}^1 = \lambda \mathbf{x}^1,$$

where λ is some number,

From the first equality of (5), taking into account that $|\mathbf{x}^1| = 1$, it follows:

$$\lambda_1 = (\lambda \mathbf{x}^1, \mathbf{x}^1) = \lambda.$$

Thus, we have proved that the maximum of quadratic form (4') on the unit sphere $|\mathbf{x}| = 1$ is reached at a certain point $\mathbf{x}^1$,

$$\max_{|\mathbf{x}|=1} (A\mathbf{x}, \mathbf{x}) = (A\mathbf{x}^1, \mathbf{x}^1) = \lambda_1.$$

And

$$A\mathbf{x}^1 = \lambda_1 \mathbf{x}^1, \quad |\mathbf{x}^1| = 1.$$

We see that a nontrivial (nonzero) vector $\mathbf{x}^1$ is mapped with the aid of operator A into $\lambda_1 \mathbf{x}^1$, i.e. into a vector collinear with $\mathbf{x}^1$.

Such vector is called the *eigenvector of operator A* , and the number λ_1 is known as the *eigenvalue belonging to this vector*.

We shall now consider the operator A on the subspace R^1 defined as the set of vectors $\mathbf{x}$ ($\in R_n$) orthogonal to the vector $\mathbf{x}^1$ (above it was denoted by L'). R^1 is an $(n - 1)$ -dimensional subspace with orthonormal bases consisting of $n - 1$ vectors. Our aim consists in finding one such basis, as we shall see, naturally related with operator A .

It is important to underline that the image $A(R^1)$ of the subspace R^1 belongs, with the aid of operator A , to R^1 , since, if $(\mathbf{x}, \mathbf{x}^1) = 0$, then

$$(A\mathbf{x}, \mathbf{x}^1) = (\mathbf{x}, A\mathbf{x}^1) = (\mathbf{x}, \lambda_1 \mathbf{x}^1) = \lambda_1 (\mathbf{x}, \mathbf{x}^1) = 0,$$

i.e. $A\mathbf{x} \in R^1$.

Self-adjointness of the operator A in R^1 is retained in a trivial way, since the equality

$$(A\mathbf{x}, \mathbf{y}) = (\mathbf{x}, A\mathbf{y})$$

being true for all $\mathbf{x}, \mathbf{y} \in R_n$, is also true for all $\mathbf{x}, \mathbf{y} \in R^1$.

And so, we now consider a self-adjoint linear operator A in a linear subspace R^1 of dimension $n - 1$. Treating it in the above way, we can reveal in R^1 a unit vector $\mathbf{x}^2$

such that

$$\max_{\substack{|\mathbf{x}|=1 \\ (\mathbf{x}, \mathbf{x}^1)=0}} (A\mathbf{x}, \mathbf{x}) = (A\mathbf{x}^2, \mathbf{x}^2) = \lambda_2 \leq \lambda_1.$$

The point is that a unit sphere S^1 in R^1 is defined, obviously, as the set of unit vectors $\mathbf{x}$ orthogonal to $\mathbf{x}^1$. In this case we may write

$$A\mathbf{x}^2 = \lambda_2\mathbf{x}^2.$$

We have found a second eigenvector of the operator A —the vector $\mathbf{x}^2$ and the eigenvalue λ_2 , belonging to it, which does not exceed λ_1 (a diminished domain of consideration may result only in a reduced maximum). And here $(\mathbf{x}^1, \mathbf{x}^2) = 0$.

Proceeding in a similar way, we can introduce a subspace R^2 of dimension $n - 2$, orthogonal to the vectors $\mathbf{x}^1$ and $\mathbf{x}^2$, show that the operator A maps R^2 into R^2 , and define a third unit vector $\mathbf{x}^3$, orthogonal to $\mathbf{x}^1$ and $\mathbf{x}^2$, such that

$$\max_{\substack{|\mathbf{x}|=1 \\ (\mathbf{x}, \mathbf{x}^1)=0, (\mathbf{x}, \mathbf{x}^2)=0}} (A\mathbf{x}, \mathbf{x}) = (A\mathbf{x}^3, \mathbf{x}^3) = \lambda_3$$

and

$$A\mathbf{x}^3 = \lambda_3\mathbf{x}^3 \quad (\lambda_3 \leq \lambda_2 \leq \lambda_1).$$

Continuing this process by induction till the n th vector $\mathbf{x}^n$, we shall obtain an orthonormal system of vectors

$$\mathbf{x}^1, \mathbf{x}^2, \dots, \mathbf{x}^n \quad (6)$$

and a system of real numbers

$$\lambda_1, \lambda_2, \dots, \lambda_n, \quad (7)$$

possessing the following properties:

$$\left. \begin{aligned} A\mathbf{x}^k &= \lambda_k\mathbf{x}^k \quad (k=1, \dots, n), \\ \lambda_1 &\geq \lambda_2 \geq \dots \geq \lambda_n. \end{aligned} \right\} \quad (8)$$

Remark 1. If a self-adjoint operator A is considered in a complex space R_n , it is possible to prove that the eigenvalues of this operator are all the same real numbers. Indeed, let $\mathbf{x}$ be the eigenvector of the self-adjoint operator

rator A and λ its corresponding eigenvalue, i.e. $A\mathbf{x} = \lambda\mathbf{x}$, $\mathbf{x} \neq 0$. Then, since A is a self-adjoint operator, $(A\mathbf{x}, \mathbf{x}) = (\mathbf{x}, A\mathbf{x})$ or $(\lambda\mathbf{x}, \mathbf{x}) = (\mathbf{x}, \lambda\mathbf{x})$.

Using the properties of the scalar product in the complex space R_n (see Sec. 6), we have: $\lambda (\mathbf{x}, \mathbf{x}) = \bar{\lambda} (\mathbf{x}, \mathbf{x})$. Since $(\mathbf{x}, \mathbf{x}) \neq 0$, we get $\lambda = \bar{\lambda}$, i.e. λ is a real number.

Besides, the eigenvectors of the self-adjoint operator A associated with different eigenvalues:

$$A\mathbf{x}_1 = \lambda_1\mathbf{x}_1, \quad A\mathbf{x}_2 = \lambda_2\mathbf{x}_2 \quad (\lambda_2 \neq \lambda_1);$$

$$(A\mathbf{x}_1, \mathbf{x}_2) = (\lambda_1\mathbf{x}_1, \mathbf{x}_2) = \lambda_1 (\mathbf{x}_1, \mathbf{x}_2);$$

$$(\mathbf{x}_1, A\mathbf{x}_2) = (\mathbf{x}_1, \lambda_2\mathbf{x}_2) = \lambda_2 (\mathbf{x}_1, \mathbf{x}_2);$$

$$0 = (\lambda_1 - \lambda_2) (\mathbf{x}_1, \mathbf{x}_2);$$

are orthogonal. Since $\lambda_1 \neq \lambda_2$, then $(\mathbf{x}_1, \mathbf{x}_2) = 0$, i.e. the eigenvectors $\mathbf{x}_1$ and $\mathbf{x}_2$ are orthogonal.

We have obtained a *complete system of eigenvectors of operator A* associated with the corresponding eigenvalues. Since orthonormal system (6) belongs to R_n and consists of n vectors, it is a basis in R_n (see Sec. 17). Therefore an arbitrary vector $\mathbf{x} \in R_n$ can be decomposed according to this system:

$$\mathbf{x} = \sum_{k=1}^n (\mathbf{x}, \mathbf{x}^k) \mathbf{x}^k. \quad (9)$$

Then our self-adjoint operator A can be written in the following way:

$$\begin{aligned} A\mathbf{x} &= A \left(\sum_{k=1}^n (\mathbf{x}, \mathbf{x}^k) \mathbf{x}^k \right) = \sum_{k=1}^n (\mathbf{x}, \mathbf{x}^k) A\mathbf{x}^k \\ &= \sum_{k=1}^n \lambda_k (\mathbf{x}, \mathbf{x}^k) \mathbf{x}^k. \end{aligned} \quad (10)$$

Thus, we have proved the following theorem:

Theorem 1. *To a self-adjoint operator A in space R_n there corresponds an orthogonal system of vectors $\mathbf{x}^1, \dots, \mathbf{x}^n$ (basis of R_n) associated with a system of real numbers $\lambda_1, \dots, \lambda_n$ such that $A\mathbf{x}$ is represented in the form of sum (10) for any $\mathbf{x} \in R_n$.*

Quadratic form (4') is then correspondingly written as follows:

$$\begin{aligned}
 (\mathbf{x}, A\mathbf{x}) &= \left(\sum_{h=1}^n (\mathbf{x}, \mathbf{x}^h) \mathbf{x}^h, \sum_{s=1}^n \lambda_s (\mathbf{x}, \mathbf{x}^s) \mathbf{x}^s \right) \\
 &= \sum_{s=1}^n \lambda_s (\mathbf{x}, \mathbf{x}^s)^2. \quad (4'')
 \end{aligned}$$

In practice, we frequently proceed from a certain quadratic form

$$\sum_{h=1}^n \sum_{l=1}^n a_{hl} x_h x_l \quad (a_{hl} = a_{lh}). \quad (4')$$

In order to apply to it the obtained results, it is possible to determine in connection with it the linear operator

$$\mathbf{y} = A\mathbf{x},$$

defined by the equalities

$$y_i = \sum_{j=1}^n a_{ij} x_j \quad (i = 1, \dots, n).$$

By virtue of the condition $a_{hl} = a_{lh}$, this is a self-adjoint operator, therefore Theorem 1 is applicable to it. In terms of the quadratic form, Theorem 1 can be reformulated as follows.

Theorem 2. *Let there be given a quadratic form (4') in a certain n -dimensional coordinate system $(x_1, \dots, x_n)$ of space R_n with the basis vectors $\mathbf{i}^1, \dots, \mathbf{i}^n$ ($\mathbf{x} = \sum_{h=1}^n x_h \mathbf{i}^h$). There exists a rectangular coordinate system $(\xi_1, \dots, \xi_n)$ with the basis vectors $\mathbf{x}^1, \dots, \mathbf{x}^n$, forming an orthogonal basis ($\mathbf{x} = \sum_{h=1}^n \xi_h \mathbf{x}^h$) and a system of real numbers $\lambda_1, \dots, \lambda_n$ such that in this system quadratic form (4') is a sum of the squares of the coordinates ξ_s of the vector $\mathbf{x}$ multiplied respectively by the numbers λ_s :*

$$\sum_{h=1}^n \sum_{l=1}^n a_{hl} x_h x_l = \sum_{s=1}^n \lambda_s \xi_s^2. \quad (4''')$$

The transition from the left-hand member of (4''') to its right-hand side can be realized if the decompositions of the vectors $\mathbf{x}^1, \dots, \mathbf{x}^n$ in terms of the basis vectors $\mathbf{i}^1, \dots, \mathbf{i}^n$ are known. Let

$$\mathbf{x}^j = \sum_{s=1}^n \beta_{js} \mathbf{i}^s$$

(see Sec. 17, (7), where a_{js} , $\mathbf{a}^s$ must be respectively replaced by β_{js} , $\mathbf{i}^s$). Since $\mathbf{i}^1, \dots, \mathbf{i}^n$ and $\mathbf{x}^1, \dots, \mathbf{x}^n$ are orthonormalized bases in R_n , the matrix

$$\Lambda = \|\beta_{js}\|$$

is orthogonal. We regard that it is known. One and the same vector $\mathbf{x}$ can be decomposed according to two bases:

$$\mathbf{x} = \sum_{s=1}^n x_s \mathbf{i}^s = \sum_{j=1}^n \xi_j \mathbf{x}^j.$$

But then

$$\sum_{j=1}^n \xi_j \mathbf{x}^j = \sum_{j=1}^n \xi_j \sum_{s=1}^n \beta_{js} \mathbf{i}^s = \sum_{s=1}^n \left(\sum_{j=1}^n \beta_{js} \xi_j \right) \mathbf{i}^s$$

and, by virtue of linear independence of the system $\mathbf{i}^1, \dots, \mathbf{i}^n$, we get

$$x_s = \sum_{j=1}^n \beta_{js} \xi_j \quad (s=1, \dots, n). \quad (11)$$

Thus, the passage from the coordinates $\xi_1, \dots, \xi_n$ to the coordinates $x_1, \dots, x_n$ is realized by means of matrix Λ^* conjugate to Λ (i.e. with the aid of the rows of Λ^* or columns of Λ).

Substituting expressions (11) for x_s into the left-hand member of (4'''), we must obtain its right-hand member. Let us write this equality:

$$\begin{aligned} \sum_{k=1}^n \sum_{l=1}^n a_{kl} \sum_{j=1}^n \beta_{jk} \xi_j \sum_{i=1}^n \beta_{il} \xi_i &= \sum_{j=1}^n \sum_{i=1}^n \sum_{k=1}^n \beta_{jk} \left(\sum_{l=1}^n a_{kl} \beta_{il} \right) \xi_j \xi_i \\ &= \sum_{s=1}^n \lambda_s \xi_s^2 = \sum_{j=1}^n \sum_{i=1}^n \delta_{ji} \lambda_i \xi_j \xi_i, \end{aligned}$$

where $\delta_{ji} = \begin{cases} 0, & j \neq i, \\ 1, & j = i, \end{cases}$ is Kronecker's symbol,

Equating the coefficients of identical $\xi_j \xi_i$, we get the following equalities

$$\sum_{k=1}^n \beta_{jk} \left(\sum_{l=1}^n a_{kl} \beta_{il} \right) = \delta_{ji} \lambda_i \quad (i, j = 1, \dots, n),$$

which can be interpreted in the following way (see Sec. 15, (6)). For the matrix

$$A = \| a_{kl} \| \quad (a_{kl} = a_{lk})$$

of the self-adjoint operator A there exists an orthogonal matrix

$$\Lambda = \| \beta_{js} \|^$$

such that

$$\Lambda \cdot A \cdot \Lambda^{-1} = \mathcal{U}, \quad (12)$$

where $\mathcal{U}$ is a certain diagonal matrix

$$\mathcal{U} = \left\| \begin{array}{cccc} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \dots & \lambda_n \end{array} \right\| \quad (13)$$

(λ_i real numbers) called *canonical*.

Note that for the orthogonal matrix Λ

$$\Lambda^{-1} = \Lambda^*.$$

Since the determinants of orthogonal matrices $|\Lambda| = |\Lambda^{-1}| = \pm 1$, it follows from (12) that

$$\prod_{j=1}^n \lambda_j = |\mathcal{U}| = |\Lambda| \cdot |A| \cdot |\Lambda^{-1}| = |A|. \quad (14)$$

We have proved, in particular, the following theorem.

Theorem 3. *If the determinant $|A|$ of a self-adjoint matrix A is not zero ($|A| \neq 0$), then all of its eigenvalues $\lambda_1, \dots, \lambda_n$ are not equal to zero ($\lambda_j \neq 0, j = 1, \dots, n$).*

It follows from Theorem 2 that

(1) If $\lambda_1 \geq \dots \geq \lambda_n > 0$, then the quadratic form is positive for any vectors $\xi \neq 0$, and, consequently, for any vectors $x \neq 0$. In this case it is called *strictly positive*,

(2) If $0 > \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$, then the form is negative for any $\xi \neq 0$, hence, for any $x \neq 0$, and is termed *strictly negative*.

(3) If $\lambda_1 \geq \dots \geq \lambda_n$ and $\lambda_n = 0$, then the form is nonnegative. There exists a direction (axis ξ_n), along which it is equal to zero. This is a *positive form*, but not *strictly*.

(4) If $\lambda_1 = 0 \geq \lambda_2 \geq \dots \geq \lambda_n$, then the form is *negative*, but not *strictly*.

(5) If $\lambda_1 > 0$, and $\lambda_n < 0$, the form is *indefinite*. With the zero point eliminated, it is positive along the axis ξ_1 and negative along the axis ξ_n .

It turns out that judging by the form of matrix $\|A\|$ and by the sign of some determinants generated by it, it is possible to find out whether its eigenvalues are all positive, all negative, or both positive and negative. This is just the essence of *Sylvester's theorem**.

Let us form a number of principal minors of the quadratic form (Ax, x) :

$$\Delta_1 = a_{11}, \quad \Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \quad \dots, \quad \Delta_n = \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{vmatrix}.$$

According to Sylvester's theorem, the following statements are possible:

1. If $\Delta_1 > 0, \Delta_2 > 0, \dots, \Delta_n > 0$, then the form is strictly positive (Case 1).

2. If $\Delta_1 < 0, \Delta_2 > 0, \Delta_3 < 0, \dots, (-1)^n \Delta_n > 0$, then the form is strictly negative (Case 2).

3. If $\Delta_1 \geq 0, \Delta_2 \geq 0, \dots, \Delta_n \geq 0$ or $\Delta_1 \leq 0, \Delta_2 \geq 0, \dots, (-1)^n \Delta_n \geq 0$ and there is j for which $\Delta_j = 0$, then the form is automatically not strictly definite.

4. In all other cases the quadratic form is indefinite.

* Sylvester, James Joseph (1814-1897). English algebraist, combinatorist, geometer, number theorist, and poet. Cofounder with Cayley of the theory of algebraic invariants (anticipated to some extent by Boole and Lagrange).

Sec. 23. Quadratic Form in Two-dimensional Space

For $n = 2$ the quadratic form has the form

$$\begin{aligned} a_{11}x_1^2 + a_{12}x_1x_2 + a_{21}x_2x_1 + a_{22}x_2^2 \\ = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2, \end{aligned} \quad (1)$$

since $a_{12} = a_{21}$.

In order to reduce form (1) to a sum of the squares of the coordinates of the vector (ξ_1, ξ_2) in a certain basis $(\mathbf{x}^1, \mathbf{x}^2)$, we have (see the preceding section) to find the basis vectors $\mathbf{x}^1, \mathbf{x}^2$ which are the eigenvectors of the self-adjoint operator A generated by the symmetric matrix

$$A = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}.$$

Let us indicate another method for finding the eigenvalues and eigenvectors of operator A , which is different from the method discussed in the preceding section.

Thus, if λ_0 is an eigenvalue of operator A and $\mathbf{x}^0 = (x_1^{(0)}, x_2^{(0)}) \neq 0$ is an eigenvector corresponding to this eigenvalue, then

$$A\mathbf{x}^0 = \lambda_0\mathbf{x}^0.$$

Let us rewrite this equation in the coordinate form:

$$\left. \begin{aligned} (a_{11} - \lambda_0)x_1^{(0)} + a_{12}x_2^{(0)} &= 0, \\ a_{21}x_1^{(0)} + (a_{22} - \lambda_0)x_2^{(0)} &= 0, \end{aligned} \right\} \quad (2)$$

or in the operator form

$$(A - \lambda_0 E)\mathbf{x}^0 = 0, \quad (2')$$

where E is an identical operator.

Thus, homogeneous system (2) has a nonzero solution $\mathbf{x}^0$ which is possible if the determinant of system (2) or (2') is zero:

$$\begin{vmatrix} a_{11} - \lambda_0 & a_{12} \\ a_{12} & a_{22} - \lambda_0 \end{vmatrix} = |A - \lambda_0 E| = 0.$$

Hence, the eigenvalue λ_0 is the root of the equation

$$|A - \lambda_0 E| = 0, \quad (3)$$

which is called the *characteristic equation* of the operator A (or of the quadratic form (Ax, x)).

The converse is also true. If λ_0 is a root of equation (3), then the nontrivial solution of the system

$$(A - \lambda_0 E) x = 0 \quad (4)$$

will be the eigenvector of the self-adjoint operator A .

Consequently, the eigenvalues of the operator A are found in this case as the roots of quadratic equation (3)

$$(a_{11} - \lambda)(a_{22} - \lambda) - a_{12}^2 = 0,$$

$$\lambda^2 - (a_{11} + a_{22})\lambda + a_{11}a_{22} - a_{12}^2 = 0.$$

Solving this equation, we get

$$\left. \begin{aligned} \lambda_1 &= \frac{1}{2} [a_{11} + a_{22} + \sqrt{4a_{12}^2 + (a_{11} - a_{22})^2}], \\ \lambda_2 &= \frac{1}{2} [a_{11} + a_{22} - \sqrt{4a_{12}^2 + (a_{11} - a_{22})^2}], \end{aligned} \right\} \quad (5)$$

whence it is seen that $\lambda_1 \geq \lambda_2$, the equality $\lambda_1 = \lambda_2$ being obtained for $a_{12} = 0$, $a_{11} = a_{22}$. For the sake of definiteness, we shall assume that $a_{11} \geq a_{22}$ (otherwise we change x_1 for x_2 and x_2 for x_1). Then $\lambda_1 \geq a_{11}$ ($\lambda_1 - a_{11} = \frac{1}{2} [\sqrt{4a_{12}^2 + (a_{11} - a_{22})^2} - (a_{11} - a_{22})] \geq 0$).

It follows from (5) that the *eigenvalues of the self-adjoint operator A are real numbers*.

Now, knowing the eigenvalues λ_1 and λ_2 , we find the unit eigenvectors as the solutions of system (4). Since $|A - \lambda_1 E| = 0$,

$$\text{rank}(A - \lambda_1 E) \leq 1.$$

If $\lambda_1 = \lambda_2$, then the matrix $A - \lambda_1 E$ consists only of zeros ($\lambda_1 = \lambda_2 = a_{11} = a_{22}$, $a_{12} = 0$), i.e. its rank is equal to zero. In this case the quadratic form is already reduced to a sum of squares ($a_{12} = a_{21} = 0$). System (4) satisfies any vector $x = (x_1, x_2)$. Therefore, the basis vectors of the coordinate system $x^1 = i = (1, 0)$, $x^2 = j = (0, 1)$ may be taken as eigenvectors. Any other system (x^1, x^2) of orthonormal vectors possesses the following property: in this system the quadratic form consists, as before, only of squares.

Now, if $\lambda_1 > \lambda_2$, then either $a_{12} \neq 0$, or $a_{12} = 0$, $a_{11} \neq a_{22}$. The second case may be omitted, since form (1) is already reduced to a sum of squares. Hence, let $a_{12} \neq 0$. Then

$$\text{rank}(A - \lambda_1 E) = 1.$$

Therefore, it is sufficient to consider one equation of system (4):

$$(a_{11} - \lambda_1) x_1 + a_{12} x_2 = 0.$$

Hence we have ($a_{12} \neq 0$)

$$x_2 = [(-a_{11} + \lambda_1)/a_{12}] x_1.$$

The vector

$$\mathbf{y}^1 = \left(x_1, \frac{\lambda_1 - a_{11}}{a_{12}} x_1 \right)$$

is the solution of system (4). Normalizing this vector, we get an eigenvector

$$\begin{aligned} \mathbf{x}^1 &= (x_1^{(1)}, x_2^{(1)}) = \frac{\mathbf{y}^1}{|\mathbf{y}^1|} \\ &= \left(\frac{\pm 1}{\sqrt{1 + \left(\frac{\lambda_1 - a_{11}}{a_{12}} \right)^2}}, \frac{\pm (\lambda_1 - a_{11})}{a_{12} \sqrt{1 + \left(\frac{\lambda_1 - a_{11}}{a_{12}} \right)^2}} \right). \end{aligned}$$

By carrying out elementary transformations, it is possible to obtain the following equalities:

$$\left. \begin{aligned} x_1^{(1)} &= \pm \frac{\text{sign } a_{12}}{\sqrt{2}} \\ &\quad \times \sqrt{1 + (a_{11} - a_{22})/\sqrt{4a_{12}^2 + (a_{11} - a_{22})^2}}, \\ x_2^{(1)} &= \pm \frac{1}{\sqrt{2}} \\ &\quad \times \sqrt{1 - (a_{11} - a_{22})/\sqrt{4a_{12}^2 + (a_{11} - a_{22})^2}}. \end{aligned} \right\} \quad (6)$$

For further purposes it is sufficient to take the plus sign in formulas (6).

Analogously, knowing the eigenvalue λ_2 , we shall find the eigenvector $\mathbf{x}^2$. It turns out that

$$\mathbf{x}^2 = (x_1^{(2)}, x_2^{(2)}) = (-x_1^{(1)}, x_2^{(1)}).$$

Let us now form the matrix of the operator (of an orthogonal transformation) Λ transferring the basis vectors $(\mathbf{i}, \mathbf{j})$ into the basis vectors $(\mathbf{x}^1, \mathbf{x}^2)$

$$\Lambda = \begin{vmatrix} x_1^{(1)} & x_2^{(1)} \\ -x_2^{(1)} & x_1^{(1)} \end{vmatrix}$$

(the rows are formed by the coordinates of the images of the basis vectors $\mathbf{i}$ and $\mathbf{j}$ with the aid of Λ , i.e. $\mathbf{x}^1 = x_1^{(1)}\mathbf{i} + x_2^{(1)}\mathbf{j}$, $\mathbf{x}^2 = -x_2^{(1)}\mathbf{i} + x_1^{(1)}\mathbf{j}$). Then the coordinates of the vector (x_1, x_2) in the system $(\mathbf{i}, \mathbf{j})$ are related with the coordinates (ξ_1, ξ_2) of this vector in the system $(\mathbf{x}^1, \mathbf{x}^2)$ with the aid of the columns of the matrix Λ :

$$\begin{cases} x_1 = x_1^{(1)}\xi_1 - x_2^{(1)}\xi_2, \\ x_2 = x_2^{(1)}\xi_1 + x_1^{(1)}\xi_2. \end{cases} \quad (7)$$

Substituting these values into quadratic form (1) and taking into consideration formulas (5) and (6), we get

$$a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 = \lambda_1\xi_1^2 + \lambda_2\xi_2^2. \quad (8)$$

The right-hand member of this equality is called the *canonical form of the quadratic form*.

If λ_1 and λ_2 are of the same sign, then the quadratic form is said to belong to the *elliptic type*; if λ_1 and λ_2 have opposite signs, then it belongs to the *hyperbolic type*; and if one of the numbers (λ_1 or λ_2) is equal to zero, then the quadratic form belongs to the *parabolic type*.

It is seen from (5) that $\lambda_1\lambda_2 = a_{11}a_{22} - a_{12}^2$. Therefore the type of form (1) can be determined by the sign of the expression $a_{11}a_{22} - a_{12}^2$.

The quadratic form will be *elliptic*, *hyperbolic*, or *parabolic* if the expression $a_{11}a_{22} - a_{12}^2$ is greater, less than, or equal to zero.

Example 1. Reduce the following to the canonical form:

$$x_1^2 - \sqrt{3}x_1x_2 + 2x_2^2.$$

In this case $a_{11} = 1$, $a_{12} = -\frac{\sqrt{3}}{2}$, $a_{22} = 2$. Since $a_{11}a_{22} - a_{12}^2 = 2 - \frac{3}{4} = \frac{5}{4} > 0$, the given quadratic form will be elliptic. We find the eigenvectors and asso-

ciated with them eigenvalues by formulas (5), (6):

$$x_1^{(1)} = \frac{1}{2}, \quad x_2^{(1)} = -\frac{\sqrt{3}}{2}, \quad \mathbf{x}^1 = \left(\frac{1}{2}, -\frac{\sqrt{3}}{2} \right), \\ \lambda_1 = \frac{5}{2}.$$

Further,

$$\mathbf{x}^2 = \left(-\frac{\sqrt{3}}{2}, \frac{1}{2} \right), \quad \lambda_2 = \frac{1}{2}.$$

In the system $(\mathbf{x}^1, \mathbf{x}^2)$ our quadratic form has the form

$$\frac{5}{2} \xi_1^2 + \frac{1}{2} \xi_2^2.$$

Since $x_1^{(1)} = \frac{1}{2} = \cos \left(-\frac{\pi}{3} \right)$, $x_2^{(1)} = -\frac{\sqrt{3}}{2} = \sin \left(-\frac{\pi}{3} \right)$, the transformation with the aid of the matrix

$$\Lambda = \begin{vmatrix} x_1^{(1)} & -x_2^{(1)} \\ x_2^{(1)} & x_1^{(1)} \end{vmatrix} \quad (\mathbf{x}^1 = (x_1^{(1)}, x_2^{(1)}), \mathbf{x}^2 = (-x_2^{(1)}, x_1^{(1)}))$$

means a rotation of the system (x_1, x_2) about the origin through an angle of $\alpha = \pi/3$ clockwise (see Example 1 in Sec. 16).

Sec. 24. Second-order Curves

In some rectangular coordinate system (x, y) , let there be given a plane curve defined implicitly by a second-degree equation

$$Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0, \quad (1)$$

where A, B, C, D, E, F are preassigned real numbers; A, B, C being not all zero. This curve is known as a *second-order* (or quadric, or quadratic) *curve*. However, it may happen that there is not a single point (x, y) with real coordinates satisfying equation (1). In such a case equation (1) is said to define an *imaginary curve of the second order*. We shall not study imaginary curves. The

equation

$$x^2 + y^2 = -1$$

may serve as an example of a second-degree equation defining an imaginary curve.

Listed below are most important particular cases of the general equation (1):

(1) The equation of an *ellipse*

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \quad (a \geq b > 0)$$

with the semiaxes of length a and b . In particular, for $a = b$ it becomes the equation of a *circle*

$$x^2 + y^2 = a^2$$

with centre at the origin and radius a .

(2) The equation of a *hyperbola*

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 \quad (a \geq b > 0)$$

with the semiaxes a and b .

(3) The equation of a *parabola*

$$y^2 = 2px \quad (p > 0).$$

(4) The equation of a *pair of intersecting straight lines*

$$a^2x^2 - b^2y^2 = 0 \quad (0 < a, b).$$

(5) The equation of a *pair of parallel or coincident straight lines*

$$x^2 - a^2 = 0 \quad (a \geq 0).$$

(6) The equation defining a *point*

$$x^2 + y^2 = 0.$$

Let us dwell briefly on the above listed curves.

The *ellipse*

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \quad (a \geq b > 0). \quad (2)$$

For $a = b$ ellipse (2) turns into a circle of radius a with centre at the origin, i.e. into a plane curve consisting of all points at a given distance a , called the ra-

dus, from a fixed point in the plane, called the centre; or, in other words: into the locus of points at the distance a from the origin.

Let $a > b$. We set $c = \sqrt{a^2 - b^2}$. Mark on the x -axis points F_1, F_2 having the abscissas $x = -c$ and $x = c$,

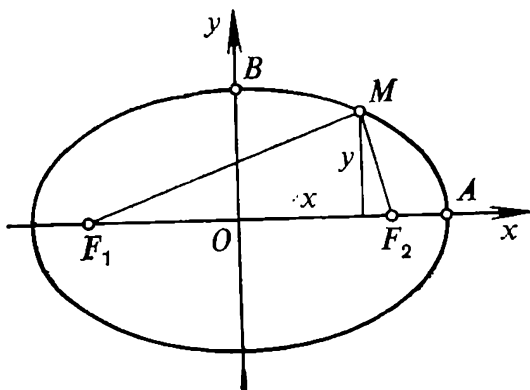


Fig. 36.

respectively. These are the *foci of the ellipse*. Ellipse (2) can be defined as the locus of points, the sum of the distances of which from the foci F_1, F_2 is a constant equal to $2a$.

Indeed (see Fig. 36),

$$MF_1 = \sqrt{(x+c)^2 + y^2}, \quad MF_2 = \sqrt{(x-c)^2 + y^2},$$

$$2a = \sqrt{(x+c)^2 + y^2} + \sqrt{(x-c)^2 + y^2},$$

whence

$$2a - \sqrt{(x+c)^2 + y^2} = \sqrt{(x-c)^2 + y^2}$$

and

$$4a^2 + (x+c)^2 + y^2 - 4a\sqrt{(x+c)^2 + y^2} = (x-c)^2 + y^2,$$

$$4a^2 + 4cx = 4a\sqrt{(x+c)^2 + y^2},$$

$$a^4 + 2a^2cx + c^2x^2 = a^2[x^2 + 2cx + c^2 + y^2],$$

$$-b^2x^2 = -a^2b^2 + a^2y^2,$$

$$a^2b^2 = b^2x^2 + a^2y^2,$$

whence equation (2) follows. Following the above computations in a reverse order, we obtain that if the point (x, y) satisfies equation (2), then the sum of its distances from F_1 and F_2 is equal to $2a$.

If we replace x by $-x$ in equation (2), then the latter will remain unchanged. This shows that ellipse (2) is a curve symmetric about the y -axis. Analogously, ellipse (2) is symmetric about the x -axis, since its equation remains unchanged when y is replaced by $-y$. But then it is sufficient to study its equation in the first quadrant (of the coordinate system), that is for $x, y \geq 0$. The portion of the ellipse found in the first quadrant is defined by the equation

$$y = \frac{b}{a} \sqrt{a^2 - x^2}, \quad 0 \leq x \leq a.$$

We see from this equation that our ellipse passes through the points $(0, b)$ and $(a, 0)$. Its ordinate y continuously decreases as x continuously increases on the interval $[0, a]$.

The ellipse is a bounded curve. It is situated inside a circle of radius a with centre at the origin (for the coordinates of any point (x, y) of the ellipse the inequality $x^2 + y^2 \leq a^2 \left(\frac{x^2}{a^2} + \frac{y^2}{b^2} \right) = a^2$ holds true).

From Fig. 36 we see that the ellipse is a continuous closed curve. In the first quadrant it is convex upward. A tangent to this curve can be drawn at any of its points*. All these and many other properties can be successfully studied using the methods of mathematical analysis which, furthermore, provides the means for exact definitions of the above mentioned notions, viz. continuity, convexity, etc.

The equation of an ellipse can also be written in the parametric form:

$$\left. \begin{aligned} x &= a \cos \theta, \\ y &= b \sin \theta, \end{aligned} \right\} \quad (-\infty < \theta < \infty). \quad (3)$$

Indeed,

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = \cos^2 \theta + \sin^2 \theta = 1,$$

* See our book "Differential and Integral Calculus", Sec. 4.2.

i.e. the point (x, y) defined by equalities (3) belongs to ellipse (2) for any θ . If θ runs continuously through the half-interval $[0, 2\pi)$, then the point (x, y) describes a complete ellipse. With a further increase in θ the motion is periodically repeated.

Let us clarify the meaning of θ and indicate the method for constructing an ellipse (Fig. 37). Draw two concentric circles of radii b and a ($b < a$) with point O as centre. Then draw a radius vector at an angle θ to the x -axis to

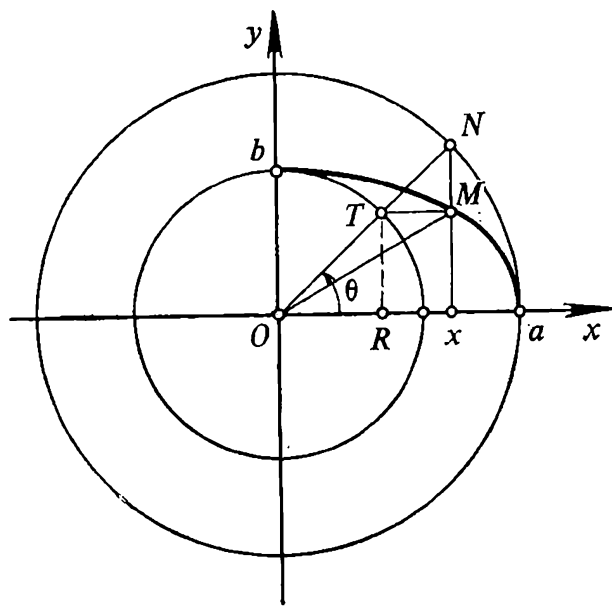


Fig. 37.

intersect the concentric circles of radii b and a at T and N , respectively. Through the point N draw a straight line parallel to the y -axis, and through the point T a straight line parallel to the x -axis. The point of intersection of these lines belongs to the ellipse. Indeed, let x be the abscissa of the point M , and y the ordinate of this point. Then (see Fig. 37)

$$x = ON \cdot \cos \theta = a \cos \theta,$$

$$y = TR = OT \cdot \sin \theta = b \sin \theta,$$

i.e. the point M is really situated in ellipse (3) and the parameter θ is the angle between the x -axis and the ray ON . Note that θ is not the polar angle φ which is formed by the radius vector OM with the axis x ($\tan \varphi = \frac{b}{a} \tan \theta$). For instance, if $\varphi = \pi/4$, $a = \sqrt{3}$, $b = 1$, then $\theta = \pi/3$; if $\varphi = 0$, then $\theta = 0$; if $\varphi = \pi/2$, then $\theta = \pi/2$.

The *hyperbola*

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 \quad (0 < a, b). \quad (4)$$

Let us set $c = \sqrt{a^2 + b^2}$ and mark on the x -axis points F_1

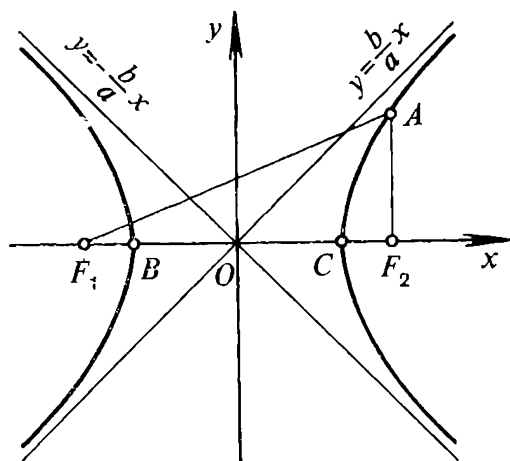


Fig. 38.

and F_2 (foci of hyperbola (4)) having the abscissas $x = -c$ and $x = c$, respectively (Fig. 38).

Hyperbola (4) can also be defined as the locus of points $A = (x, y)$, whose distances from the foci F_1 and F_2 have a constant difference equal to $2a$.

We have (see Fig. 38)

$$AF_1 - AF_2 = \sqrt{(x+c)^2 + y^2} - \sqrt{(x-c)^2 + y^2} = 2a,$$

$$\sqrt{(x+c)^2 + y^2} = \sqrt{(x-c)^2 + y^2} + 2a,$$

$$(x+c)^2 + y^2 = (x-c)^2 + y^2 + 4a \sqrt{(x-c)^2 + y^2} + 4a^2,$$

$$4cx - 4a^2 = 4a \sqrt{(x-c)^2 + y^2},$$

$$\begin{aligned}
c^2x^2 - 2a^2cx + a^4 &= a^2(x^2 - 2cx + c^2) + a^2y^2, \\
(a^2 + b^2)x^2 &= a^2x^2 + a^2b^2 + a^2y^2, \\
b^2x^2 - a^2y^2 &= a^2b^2,
\end{aligned}$$

whence equation (4) follows.

We have obtained the right-hand branch of hyperbola (4). In order to get its left-hand branch, we have to proceed from the equality

$$AF_2 - AF_1 = 2a.$$

Reasoning in a reverse order and proceeding from equation (4), we can conclude that the points (x, y) satisfying this equation belong to the above indicated locus.

Judging by the form of equation (4), we conclude that hyperbola (4) is symmetric about the x - and y -axes. The portion of hyperbola (4) located in the first quadrant has the equation

$$y = \frac{b}{a} \sqrt{x^2 - a^2} \quad (a \leq x < \infty). \quad (5)$$

We see that our hyperbola passes through the point $(a, 0)$, and with an increase in x on the half-interval $[a, \infty)$ the ordinate y increases and tends to infinity. The points $B = (-a, 0)$ and $C = (a, 0)$ at which the hyperbola intersects the x -axis are called the *vertices of the hyperbola*.

Figure 38 shows two straight lines:

$$y = \pm \frac{b}{a} x.$$

These are *asymptotes* of our hyperbola.

Let on the half-interval $[a, \infty)$ (or $(-\infty, a]$) there be given a curve $y = f(x)$. The straight line $y = mx + n$ is said to be an *asymptote* of this curve as $x \rightarrow +\infty$ ($x \rightarrow -\infty$) if

$$\lim_{x \rightarrow +\infty} [f(x) - mx - n] = 0$$

(respectively, $\lim_{x \rightarrow -\infty} [f(x) - mx - n] = 0$).

Let us consider the piece of our hyperbola defined by equality (5) and compare it with the straight line $y =$

$= \frac{b}{a} x$. The limit

$$\lim_{x \rightarrow +\infty} \left[\frac{bx}{a} - \frac{b}{a} \sqrt{x^2 - a^2} \right] = \lim_{x \rightarrow +\infty} \frac{b}{a} \frac{a^2}{x + \sqrt{x^2 - a^2}} = 0.$$

This shows that the straight line $y = \frac{b}{a} x$ is the asymptote of the piece under consideration as $x \rightarrow +\infty$. But then the straight line $y = \frac{b}{a} x$ is said to be the asymptote of the (entire!) hyperbola as $x \rightarrow +\infty$. By virtue of symmetry of our hyperbola about the axes, and the symmetry of the pair of straight lines $y = \pm \frac{b}{a} x$ about the axes, we may say that both of these lines are asymptotes of our hyperbola as $x \rightarrow +\infty$, and as $x \rightarrow -\infty$.

The right-hand branch of hyperbola (4) can be represented parametrically

$$\left. \begin{aligned} x &= a \cosh u = \frac{a}{2} (e^u + e^{-u}), \\ y &= b \sinh u = \frac{b}{2} (e^u - e^{-u}), \end{aligned} \right\} \quad (-\infty < u < \infty). \quad (6)$$

Indeed, since

$$\cosh^2 u - \sinh^2 u = 1, \quad (7)$$

we get from equations (6)

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = \cosh^2 u - \sinh^2 u = 1.$$

The upper half of the right-hand branch of the hyperbola corresponds to the variation of the parameter $u \in [0, \infty)$, while the lower half to $u \in (-\infty, 0]$.

Let us clarify how the parameter u is related with the parameter θ in the parametric equation of the ellipse and, at the same time, indicate how to construct a hyperbola with the aid of compasses and a ruler. Since our method of constructing the hyperbola is based on the method of constructing the ellipse, we are going to present these two construction methods jointly (see Fig. 39). We shall confine ourselves to the construction of the parts of ellipse (2) and hyperbola (6) situated in the first quadrant.

Draw two concentric circles of radii a and b with the origin as centre. Draw the ray emanating from the origin at an angle θ_0 to the x -axis. Let T_1 and N_1 be the points of intersection of this ray and the concentric circles ($OT_1 = b$, $ON_1 = a$). Drawing through the points T_1 and N_1 straight lines parallel to the x - and y -axes, respectively, we shall obtain the point of their intersection

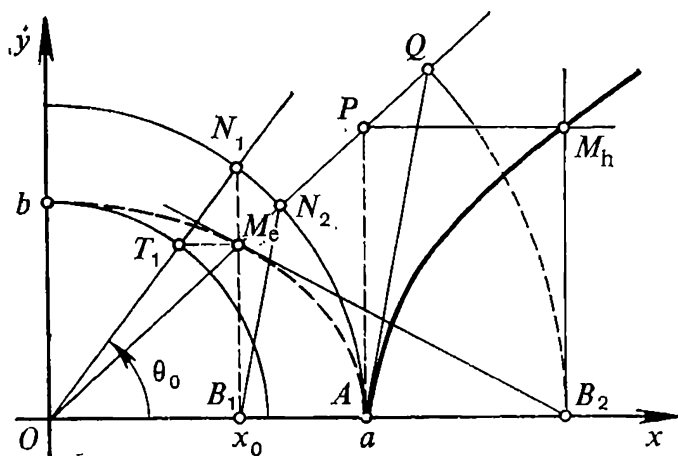


Fig. 39.

$M_e = (x_0, y_0)$ belonging to ellipse (2). We then draw the ray OM_e . Let N_2 be the point of intersection of this ray with the circle of radius a , and P the point of intersection of this ray and the straight line parallel to the y -axis and passing through the point $A = (a, 0)$ belonging to the ellipse. The equation of the ray OP can be written as follows

$$y = \frac{y_0}{x_0} x.$$

Hence it follows that the ordinate of the point P is equal to $Y_0 = \frac{ay_0}{x_0}$. Further, join the points $B_1 = (x_0, 0)$ and N_2 by a straight line, and, through the point A , draw a straight line, parallel to B_1N_2 , to intersect the ray OP at the point Q . From the similarity of the triangles OAQ and OB_1N_2 we obtain that $OQ = a^2/x_0$. Describing an

arc of radius OQ from the origin as centre, we intersect the x -axis at the point $B_2 = (a^2/x_0, 0)$.

Now we draw, through the points B_2 and P , straight lines parallel to the y - and x -axes, respectively. The point of intersection of these lines $M_h = (X_0, Y_0)$, where $X_0 = a^2/x_0$, belongs to hyperbola (4).

Indeed, since the point (x_0, y_0) lies on ellipse (2), we get

$$\frac{X_0^2}{a^2} - \frac{Y_0^2}{b^2} = \frac{a^4}{a^2 x_0^2} - \frac{a^2 y_0^2}{x_0^2 b^2} = \frac{a^2}{x_0^2} \left(1 - \frac{y_0^2}{b^2} \right) = \frac{a^2}{x_0^2} \cdot \frac{x_0^2}{a^2} = 1,$$

i.e. the point M_h belongs to hyperbola (4).

Note that $B_2 = (a^2/x_0, 0)$ is the point of intersection of the tangent to the ellipse at the point M_e with the x -axis (see the last footnote).

Thus, to every point (x, y) of ellipse (2) there corresponds a quite definite point (X, Y) of hyperbola (4), and vice versa.

Now, if ellipse (2) is represented parametrically, then

$$x = a \cos \theta, \quad y = b \sin \theta.$$

Therefore,

$$X = \frac{a^2}{x} = \frac{a}{\cos \theta}, \quad Y = \frac{ay}{x} = b \tan \theta.$$

Hence, taking into account (6), we get

$$\cosh u = \frac{1}{\cos \theta}, \quad \sinh u = \tan \theta.$$

The following formulas also hold true:

$$\begin{aligned} \tan \frac{\theta}{2} &= \sqrt{\frac{1 - \cos \theta}{1 + \cos \theta}} = \sqrt{\frac{\cosh u - 1}{\cosh u + 1}} = \tanh \frac{u}{2}; \\ e^u &= \cosh u + \sinh u = \frac{1}{\cos \theta} + \tan \theta = \frac{1 + \sin \theta}{\cos \theta} \\ &= \frac{\left(\cos \frac{\theta}{2} + \sin \frac{\theta}{2} \right)^2}{\cos^2 \frac{\theta}{2} - \sin^2 \frac{\theta}{2}} = \frac{\cos \frac{\theta}{2} + \sin \frac{\theta}{2}}{\cos \frac{\theta}{2} - \sin \frac{\theta}{2}} = \frac{\sin \left(\frac{\pi}{4} + \frac{\theta}{2} \right)}{\cos \left(\frac{\pi}{4} + \frac{\theta}{2} \right)} \\ &= \tan \left(\frac{\pi}{4} + \frac{\theta}{2} \right), \end{aligned}$$

i.e.

$$u = \ln \tan \left(\frac{\pi}{4} + \frac{\theta}{2} \right).$$

The *parabola*

$$y^2 = 2px \quad (p > 0). \quad (8)$$

Let us mark on the x -axis a point F with the abscissa $x = p/2$ called the *focus* of parabola (8) and draw a straight line $x = -p/2$ termed the *directrix* of parabola (8) (Fig. 40).

Parabola (8) can also be defined as the locus of points

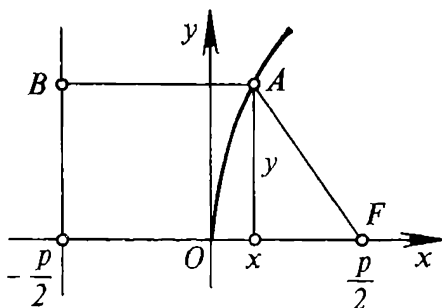


Fig. 40.

$A = (x, y)$ equidistant from the focus and directrix. Indeed (see Fig. 40)

$$AF^2 = \left(x - \frac{p}{2} \right)^2 + y^2,$$

$$AB^2 = \left(x + \frac{p}{2} \right)^2,$$

consequently,

$$\left(x - \frac{p}{2} \right)^2 + y^2 = \left(x + \frac{p}{2} \right)^2$$

$$-px + y^2 = px,$$

i.e.

$$y^2 = 2px.$$

Conversely, it follows from this equation that the points, satisfying it, belong to the indicated locus of points.

As is seen from equation (8), parabola (8) is symmetric about the x -axis. Its upper half has the equation

$$y = \sqrt{2px} \quad (0 \leq x < \infty), \quad (9)$$

wherefrom it is seen that when x increases, running through the half-interval $[0, \infty)$, the ordinate y also increases from 0 to ∞ .

Let us indicate a simple method for constructing parabola (9) with the aid of a ruler and a right angle, or with the aid of compasses and a ruler. Draw a straight line

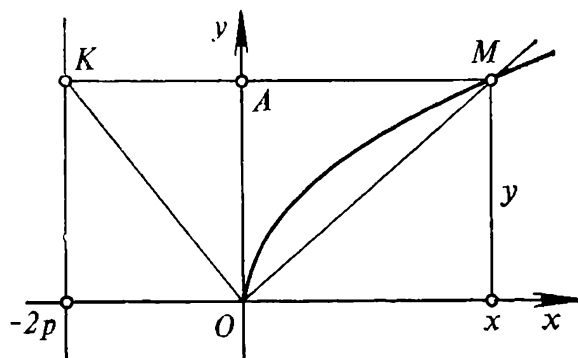


Fig. 41.

$x = -2p$ (Fig. 41) and take on it an arbitrary point $K = (-2p, y)$, $y > 0$. Join this point to the origin and draw a straight line passing through the origin perpendicular to the line OK . Draw then a straight line through the point K parallel to the x -axis. The latter two lines intersect at the point $M = (x, y)$ which belongs to parabola (9), since $OA = y$ is the geometric mean of the numbers $2p$ and x ($y = \sqrt{2px}$).

Parabola (8) has no asymptotes*.

A pair of intersecting lines

$$a^2x^2 - b^2y^2 = (ax - by)(ax + by) = 0 \quad (0 < a, b). \quad (10)$$

* See our book "Differential and Integral Calculus", Sec. 4.20.

If a certain point (x, y) satisfies equation (10), then it satisfies one of the equations

$$\left. \begin{aligned} ax - by &= 0, \\ ax + by &= 0 \end{aligned} \right\} \quad (10')$$

or both of them. Conversely, if the point (x, y) satisfies one of the equations (10'), then it also satisfies equation (10). In this meaning (10) is said to be the equation of a pair of straight lines.

We shall prove below that there exists a rectangular coordinate system such that in it curve (1), provided that it is not an imaginary one, has one of the above listed equations (1) to (6).

Namely:

for $AC - B^2 > 0$ curve (1) is an ellipse, a point (Cases (1), (6)), or an imaginary curve;

for $AC - B^2 < 0$ curve (1) is a hyperbola or a pair of intersecting (different) straight lines (Cases (2), (4));

for $AC - B^2 = 0$ curve (1) is a parabola, a pair of parallel or coincident straight lines, or an imaginary curve (Cases (3), (5)).

We allow ourselves to use "curve" even in Cases (4), (5), (6) when we speak of a pair of straight lines or of a set consisting of one point.

And so, let there be given an equation

$$Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0, \quad (1)$$

where the coefficients A, B, C are not all zero.

Without the loss of generality, we may regard that $A \geq 0, A \geq C, B \geq 0$. We can always arrive to this situation with the aid of the orthogonal transformations:

$$\left. \begin{aligned} x &= \eta, \\ y &= \xi, \end{aligned} \right\} \quad \left. \begin{aligned} x &= -\xi, \\ y &= \eta \end{aligned} \right\}$$

and by multiplying both sides of (1) by -1 .

If $B = 0, A \geq C > 0$, then (1) can be written in the form

$$A \left(x + \frac{D}{A} \right)^2 + C \left(y + \frac{E}{C} \right)^2 + F - \frac{E^2}{C} - \frac{D^2}{A} = 0. \quad (11)$$

The translation

$$\xi = x + \frac{D}{A}, \quad \eta = y + \frac{E}{C}$$

transforms equation (11) in the following way:

$$A\xi^2 + C\eta^2 = \frac{E^2}{C} + \frac{D^2}{A} - F. \quad (11')$$

If the number $\frac{E^2}{C} + \frac{D^2}{A} - F > 0$, then equation (11') represents the equation of an ellipse (Case (1)) with the semiaxes a, b , where

$$a^2 = \left(\frac{E^2}{C} + \frac{D^2}{A} - F \right) / A, \quad b^2 = \left(\frac{E^2}{C} + \frac{D^2}{A} - F \right) / C.$$

Note that in this case $AC - B^2 = AC > 0$.

And if the right-hand member of equation (11') is equal to zero, then we get a point (Case (6)).

With the negative right-hand member, equation (11') yields an imaginary curve.

If $C = 0$, $A > 0$, then equation (1) can be written in the form:

$$A \left(x + \frac{D}{A} \right)^2 + 2Ey + F - \frac{D^2}{A} = 0. \quad (12)$$

Let the number $E \neq 0$, then the translation

$$\xi = x + \frac{D}{A}, \quad \eta = y + \frac{F}{2E} - \frac{D^2}{2AE}$$

transforms (12) into the equation

$$A\xi^2 + 2E\eta = 0, \quad (12')$$

which (on replacing, if necessary, η by $-\eta$) represents the equation of a parabola (Case (3)).

If $E = 0$, then, depending on the sign of the number $\frac{D^2}{A} - F$, we shall obtain either a pair of parallel straight lines, or an imaginary curve. Note that here $AC - B^2 = 0$.

Further, if the number $C < 0$, $A > 0$, then equation (1) can be written as follows:

$$A \left(x + \frac{D}{A} \right)^2 - |C| \left(y - \frac{E}{|C|} \right)^2 + F + \frac{E^2}{|C|} - \frac{D^2}{A} = 0, \quad (13)$$

which is analyzed in the same way as equation (11). Equation (13) yields either a hyperbola, or a pair of intersecting straight lines (Cases (2) and (4)). Note that in this case $AC - B^2 = AC < 0$.

The case $A = 0, C < 0$ is reduced to an equation of type (12).

Hence, for $B = 0$ equation (1) is reduced to one of particular cases (1) to (6).

Let now $B > 0, A \geq C$. Then, as we know from the preceding section, there exists an orthogonal transformation

$$\left. \begin{aligned} x &= x_1 \xi - y_1 \eta, \\ y &= y_1 \xi + x_1 \eta, \end{aligned} \right\} \quad (14)$$

where

$$\begin{aligned} x_1 &= \sqrt{\frac{1}{2} + \frac{A-C}{2\sqrt{4B^2 + (A-C)^2}}}, \\ y_1 &= \sqrt{\frac{1}{2} - \frac{A-C}{2\sqrt{4B^2 + (A-C)^2}}}, \end{aligned}$$

which reduces the quadratic form

$$Ax^2 + 2Bxy + Cy^2$$

to the canonical form.

Let us transform equation (1) with the aid of (14):

$$\lambda_1 \xi^2 + \lambda_2 \eta^2 + 2D(x_1 \xi - y_1 \eta) + 2E(y_1 \xi + x_1 \eta) + F = 0, \quad (15)$$

where

$$\begin{aligned} \lambda_1 &= \frac{1}{2} [A + C + \sqrt{4B^2 + (A - C)^2}], \\ \lambda_2 &= \frac{1}{2} [A + C - \sqrt{4B^2 + (A - C)^2}] \end{aligned}$$

($\lambda_1 > \lambda_2, \lambda_1 \lambda_2 = AC - B^2$). We rewrite equation (15) in the form

$$\begin{aligned} \lambda_1 \xi^2 + \lambda_2 \eta^2 + 2(x_1 D + y_1 E) \xi + 2(x_1 E - y_1 D) \eta + F &= 0. \end{aligned} \quad (15')$$

Equation (15') is a particular case of equation (1) for $B = 0$, which was already analyzed.

Thus, we may state that if:

(1) $AC - B^2 = \lambda_1 \lambda_2 > 0$, then equation (1) represents either an ellipse, or a point, or an imaginary curve. In this case we shall say that equation (1) belongs to the *elliptic type*;

(2) $AC - B^2 = \lambda_1 \lambda_2 < 0$, then equation (1) represents either a hyperbola, or a pair of intersecting straight lines. In this case equation (1) will be said to belong to the *hyperbolic type*;

(3) $AC - B^2 = \lambda_1 \lambda_2 = 0$, then equation (1) represents either a parabola, or a pair of parallel lines, or an imaginary curve. In this case we shall say that equation (1) belongs to the *parabolic type*.

Example 1. Find out the character of the curve

$$2x^2 + \sqrt{3}xy + y^2 + 2x + 2\sqrt{3}y + F = 0,$$

where F is an arbitrary real number.

Here $A = 2 > C = 1$, $B = \frac{\sqrt{3}}{2} > 0$, $AC - B^2 = \frac{5}{4} > 0$, i.e. the equation belongs to the elliptic type. It is easy to compute (see the example in the preceding section) that

$$x_1 = \frac{\sqrt{3}}{2}, \quad y_1 = \frac{1}{2}, \quad \lambda_1 = \frac{5}{2}, \quad \lambda_2 = \frac{1}{2}.$$

Therefore, with the aid of the orthogonal transformation

$$x = \frac{1}{2}(\sqrt{3}\xi - \eta),$$

$$y = \frac{1}{2}(\xi + \sqrt{3}\eta)$$

our equation will be written as follows

$$\frac{5}{2}\xi^2 + \frac{1}{2}\eta^2 + (\sqrt{3}\xi - \eta) + \sqrt{3}(\xi + \sqrt{3}\eta) + F = 0$$

or

$$\frac{5}{2}\left(\xi + \frac{2}{5}\sqrt{3}\right)^2 + \frac{1}{2}(\eta + 2)^2 = \frac{16}{5} - F.$$

In addition, let us realize the translation

$$u = \xi + \frac{2}{5} \sqrt{3},$$

$$v = \eta + 2,$$

then we shall have

$$\frac{5}{2} u^2 + \frac{1}{2} v^2 = \frac{16}{5} - F. \quad (16)$$

If $\frac{16}{5} - F > 0$, then (16) will be the equation of an ellipse with the semiaxes a , b , where

$$a^2 = 2(16 - 5F)/25, \quad b^2 = 2(16 - 5F)/5.$$

If $\frac{16}{5} - F = 0$, then (16) yields a point. If $\frac{16}{5} - F < 0$, then (16) represents an imaginary curve.

Sec. 25. Second-order Surfaces in Three-dimensional Space

The equation

$$\sum_{k=1}^3 \sum_{l=1}^3 a_{kl} x_k x_l + 2 \sum_{l=1}^3 A_l x_l + B = 0, \quad (1)$$

where $a_{kl} = a_{lk}$, A_l , B are preassigned constants, and $\mathbf{x} = (x_1, x_2, x_3)$ is a variable point in R_3 defines, generally speaking, a certain set of points in R_3 called the *surface of the second order*. If equation (1) is satisfied by none of real points $\mathbf{x} = (x_1, x_2, x_3)$ then it is said to define an *imaginary surface*. However, we are not interested in these cases. In some cases equation (1) may define a pair of different or coincident planes or only one point. But such sets will also be referred to as surfaces.

Let us enumerate the most important particular cases of equation (1):

(1) *Ellipsoid*

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1 \quad (a, b, c > 0).$$

(2) *Hyperboloid of one sheet*

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1 \quad (a, b, c > 0).$$

(3) *Hyperboloid of two sheets*

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1 \quad (a, b, c > 0).$$

(4) *Elliptic paraboloid*

$$\frac{x^2}{p} + \frac{y^2}{q} = 2z \quad (p, q > 0).$$

(5) *Hyperbolic paraboloid*

$$\frac{x^2}{p} - \frac{y^2}{q} = 2z \quad (p, q > 0).$$

(6) *Cone of the second order*

$$\frac{x^2}{b^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 0 \quad (a, b, c > 0),$$

(7) *Point*

$$x^2 + y^2 + z^2 = 0.$$

(8) *Cylinders of the second order:*

elliptic cylinder

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \quad (a, b > 0),$$

hyperbolic cylinder

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 \quad (a, b > 0),$$

parabolic cylinder

$$y^2 = 2px \quad (p > 0),$$

pair of intersecting planes

$$a^2x^2 - b^2y^2 = 0 \quad (a, b > 0),$$

pair of parallel or coincident planes

$$x^2 - a^2 = 0 \quad (a > 0),$$

$$z^2 = 0,$$

straight line

$$x^2 + y^2 = 0.$$

When considering the particular cases of equation (1) we regarded that $x = x_1$, $y = x_2$, $z = x_3$.

It is possible to prove that for each particular equation (1), provided that it defines a real (not an imaginary!) surface, we can find a rectangular coordinate system in which this equation has one of the above listed eight forms. This follows from the general theory set forth in Sec. 22. The transformation of equation (1) itself is carried out in the same way as in the preceding section. The finding of the eigenvalues λ_1 , λ_2 , λ_3 is reduced to solving a cubic equation.

We would like to indicate one more way of finding eigenvalues and eigenvectors, which, as a matter of fact, was considered in Sec. 23 for the two-dimensional case. The eigenvalues λ_1 , λ_2 , λ_3 of the self-adjoint operator A and associated with them normalized vectors $\mathbf{x}^1$, $\mathbf{x}^2$, $\mathbf{x}^3$ ($|\mathbf{x}^j| = 1$, $A\mathbf{x}^j = \lambda_j\mathbf{x}^j$, $j = 1, 2, 3$) can be found in the following way (the proof is given below).

We introduce the determinant ($a_{kl} = a_{lk}$)

$$D(\lambda) = |A - \lambda E| = \begin{vmatrix} a_{11} - \lambda & a_{12} & a_{13} \\ a_{21} & a_{22} - \lambda & a_{23} \\ a_{31} & a_{32} & a_{33} - \lambda \end{vmatrix},$$

where E is a unit matrix. Find the roots λ_1 , λ_2 , λ_3 of the equation

$$|A - \lambda E| = 0, \quad (2)$$

which is called the *characteristic equation* of the operator A ($D(\lambda_j) = 0$, $j = 1, 2, 3$). These are just eigenvalues of the operator A . Being real, they may be different, but may also coincide, i.e. be multiple. Hence,

$$D(\lambda) = (\lambda_1 - \lambda)(\lambda_2 - \lambda)(\lambda_3 - \lambda).$$

Then for the root λ_1 we look for the nontrivial solution $\mathbf{x} = (x_1, x_2, x_3)$ of the homogeneous system of equations:

$$\left. \begin{aligned} (a_{11} - \lambda_1)x_1 + a_{12}x_2 + a_{13}x_3 &= 0, \\ a_{12}x_1 + (a_{22} - \lambda_1)x_2 + a_{23}x_3 &= 0, \\ a_{31}x_1 + a_{32}x_2 + (a_{33} - \lambda_1)x_3 &= 0 \end{aligned} \right\} \quad (3)$$

or of the corresponding equation for the operator $A - \lambda_1 E$:

$$(A - \lambda_1 E) \mathbf{x} = 0. \quad (3')$$

where E is a unit matrix and $\mathbf{x} = (x_1, x_2, x_3)$, $\mathbf{0} = (0, 0, 0)$ are vectors.

If λ_1 is a simple root (i.e. in this case λ_1 is different from λ_2 and from λ_3), then the rank of the matrix of system (3) will be necessarily equal to two (rank $(A - \lambda_1 E) = 2$) and we shall get the only (up to a sign) vector $\mathbf{x}^1 = (x_1^{(1)}, x_2^{(1)}, x_3^{(1)})$, satisfying system (3), i.e.

$$(A - \lambda_1 E) \mathbf{x}^1 = 0$$

$$A \mathbf{x}^1 = \lambda_1 \mathbf{x}^1.$$

If λ_1 is a root of the second multiplicity ($\lambda_1 = \lambda_2 \neq \lambda_3$), then system (3) necessarily has the rank equal to unity (rank $(A - \lambda_1 E) = 1$), and will have two orthonormal solutions $\mathbf{x}^1$ and $\mathbf{x}^2$ ($|\mathbf{x}^1| = |\mathbf{x}^2| = 1$, $(\mathbf{x}^1, \mathbf{x}^2) = 0$) which are eigenvectors associated with the eigenvalue λ_1 :

$$A \mathbf{x}^j = \lambda_j \mathbf{x}^j \quad (j = 1, 2, \lambda_1 = \lambda_2).$$

Finally, if λ_1 is a root of the third multiplicity ($\lambda_1 = \lambda_2 = \lambda_3$), then system (3) necessarily has the rank equal to zero (rank $(A - \lambda_1 E) = 0$), and will have three orthonormal solutions $\mathbf{x}^1, \mathbf{x}^2, \mathbf{x}^3$:

$$A \mathbf{x}^j = \lambda_j \mathbf{x}^j \quad (j = 1, 2, 3, \lambda_1 = \lambda_2 = \lambda_3).$$

Any three orthonormal vectors in R_3 may be taken as eigenvectors $\mathbf{x}^1, \mathbf{x}^2, \mathbf{x}^3$ belonging to the eigenvalues $\lambda_1 = \lambda_2 = \lambda_3$.

Let us prove this. We know that in R_3 there exists a system of orthonormal vectors $\mathbf{x}^1, \mathbf{x}^2, \mathbf{x}^3$ and real numbers $\lambda_1, \lambda_2, \lambda_3$ such that

$$A \mathbf{x}^j = \lambda_j \mathbf{x}^j \quad (j = 1, 2, 3),$$

And we can indicate an orthogonal matrix Λ such that (see Sec. 22)

$$\Lambda A \Lambda^{-1} = \begin{vmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{vmatrix}.$$

Hence, for a variable number λ the following identity holds true

$$\Lambda (A - \lambda E) \Lambda^{-1} = \begin{vmatrix} \lambda_1 - \lambda & 0 & 0 \\ 0 & \lambda_2 - \lambda & 0 \\ 0 & 0 & \lambda_3 - \lambda \end{vmatrix} \quad (4)$$

$$\begin{aligned} (\Lambda (A - \lambda E) \Lambda^{-1}) \Lambda^{-1} &= \Lambda A \Lambda^{-1} - \Lambda \lambda E \Lambda^{-1} = \\ &= \Lambda A \Lambda^{-1} - \lambda \Lambda \Lambda^{-1} = \Lambda A \Lambda^{-1} - \lambda E. \end{aligned}$$

Above, we denoted the determinant of the matrix $A - \lambda E$ by $D(\lambda)$. As it is seen from (4), it is equal to the determinant of the matrix standing in the right-hand side of (4), since

$$\begin{aligned} (\lambda_1 - \lambda) (\lambda_2 - \lambda) (\lambda_3 - \lambda) &= |\Lambda (A - \lambda E) \Lambda^{-1}| \\ &= |\Lambda| |A - \lambda E| |\Lambda^{-1}| = |A - \lambda E| = D(\lambda). \end{aligned}$$

Thus,

$$D(\lambda) = (\lambda_1 - \lambda) (\lambda_2 - \lambda) (\lambda_3 - \lambda).$$

We have obtained that the roots of the polynomial $D(\lambda)$ coincide with the eigenvalues $\lambda_1, \lambda_2, \lambda_3$ of the operator A . Hence, they are *real*.

Let λ_1 be a simple root and, consequently, $\lambda_1 \neq \lambda_2$ and $\lambda_1 \neq \lambda_3$. Then the matrix in the right-hand side of (4) has for $\lambda = \lambda_1$ the rank equal to two (rank $\Lambda (A - \lambda_1 E) \Lambda^{-1} = 2$), but then rank $(A - \lambda_1 E) = 2$. As a matter of fact, the solutions of homogeneous system (3) and the system

$$\left. \begin{aligned} 0 \cdot z_1 + 0 \cdot z_2 + 0 \cdot z_3 &= 0, \\ 0 \cdot z_1 + (\lambda_2 - \lambda_1) z_2 + 0 \cdot z_3 &= 0, \\ 0 \cdot z_1 + 0 \cdot z_2 + (\lambda_3 - \lambda_1) z_3 &= 0, \end{aligned} \right\} \quad (5)$$

corresponding to matrix (4) are transformed into each other with the aid of an orthogonal matrix (systems (3) and (5) are equivalent). But system (5) has only one (up to a sign) normalized solution $(\pm 1, 0, 0)$. And this is possible only if $\text{rank } (A - \lambda_1 E) = 2$.

It is possible to give the following explanation of this fact. If we assume that all second-order determinants generated by the matrix $A - \lambda_1 E$ are equal to zero, then all the second-order determinants generated by the matrix $\Lambda (A - \lambda_1 E) \Lambda^{-1}$, will also equal zero, since these determinants are linear combinations of the second-order determinants from the matrix $A - \lambda_1 E$. But this is impossible, since the determinant generated by the matrix $\Lambda (A - \lambda_1 E) \Lambda^{-1}$,

$$\begin{vmatrix} \lambda_2 - \lambda_1 & 0 \\ 0 & \lambda_3 - \lambda_1 \end{vmatrix} \neq 0.$$

If now $\lambda_1 = \lambda_2 \neq \lambda_3$, then, reasoning in a similar way, we shall obtain that the $\text{rank } (A - \lambda_1 E) = 1$ and, hence, there exist exactly two orthonormal solutions $\mathbf{x}^1, \mathbf{x}^2$ of system (3) corresponding to $\lambda_1 = \lambda_2$.

Finally, for $\lambda_1 = \lambda_2 = \lambda_3$ $\text{rank } (A - \lambda_1 E) = 0$, i.e. all the elements of the matrix $A - \lambda_1 E$ are zero. In this case any vector $\mathbf{x} = (x_1, x_2, x_3)$ is a solution of system (3). This results in that any three vectors $\mathbf{x}^1, \mathbf{x}^2, \mathbf{x}^3$ forming an orthonormal system will be the eigenvectors belonging to the eigenvalue $\lambda_1 = \lambda_2 = \lambda_3$.

Note that in this situation the quadratic form is already reduced to the sum of squares ($a_{12} = a_{13} = a_{23} = 0, a_{11} = a_{22} = a_{33} = \lambda_1$).

Example 1. Reduce the following to the canonical form

$$x_1^2 + x_2^2 + 2x_1x_2 + 2x_1x_3 + 2x_2x_3.$$

Here $a_{11} = a_{22} = a_{12} = a_{13} = a_{23} = 1, a_{33} = 0$. Let us set up a characteristic equation

$$\begin{vmatrix} 1 - \lambda & 1 & 1 \\ 1 & 1 - \lambda & 1 \\ 1 & 1 & -\lambda \end{vmatrix} = 0$$

or

$$\begin{aligned} -\lambda (1 - \lambda)^2 + 2 - 2(1 - \lambda) + \lambda &= 0, \\ -\lambda (1 - \lambda)^2 + 3\lambda &= 0. \end{aligned}$$

It is easy to see that $\lambda = 0$ is a root of this equation. Let us find two other roots:

$$\begin{aligned} 3 - (1 - \lambda)^2 &= 0, \quad (1 - \lambda)^2 = 3, \quad 1 - \lambda = \pm\sqrt{3}, \\ \lambda &= 1 \pm \sqrt{3}. \end{aligned}$$

Hence,

$$\lambda_1 = 1 + \sqrt{3}, \quad \lambda_2 = 0, \quad \lambda_3 = 1 - \sqrt{3},$$

i.e. we have obtained the case: $\lambda_1 > \lambda_2 > \lambda_3$.

Let us find the eigenvector $\mathbf{x}^1$. To this effect, we set up system (3):

$$\begin{aligned} -\sqrt{3}x_1 + x_2 + x_3 &= 0 \\ x_1 - \sqrt{3}x_2 + x_3 &= 0 \\ x_1 + x_2 - (1 + \sqrt{3})x_3 &= 0. \end{aligned}$$

Any two equations of this system are linearly independent. Solving the system of the first two equations, we get

$$x_1 = \frac{1 + \sqrt{3}}{2} x_3, \quad x_2 = \frac{1 + \sqrt{3}}{2} x_3.$$

Thus, the vector

$$\mathbf{y}^1 = \left(\frac{1 + \sqrt{3}}{2} x_3, \quad \frac{1 + \sqrt{3}}{2} x_3, x_3 \right)$$

is a solution of the system and, normalizing it, we get the eigenvector

$$\mathbf{x}^1 = \frac{\mathbf{y}^1}{|\mathbf{y}^1|} = \left(\frac{1 + \sqrt{3}}{2\sqrt{3 + \sqrt{3}}}, \quad \frac{1 + \sqrt{3}}{2\sqrt{3 + \sqrt{3}}}, \quad \frac{1}{\sqrt{3 + \sqrt{3}}} \right).$$

Let us find $\mathbf{x}^2$ ($\lambda_2 = 0$):

$$\begin{aligned} x_1 + x_2 + x_3 &= 0 \\ x_1 + x_2 + x_3 &= 0 \\ x_1 + x_2 + 0 \cdot x_3 &= 0, \end{aligned}$$

Solving the system of the last two equations (since the determinant made up of the coefficients of x_2 and x_3 is not equal to zero), we obtain $x_2 = -x_1$, $x_3 = 0$.

The vector $\mathbf{y}^2 = (x_1, -x_1, 0)$ is the solution of the system and the vector

$$\mathbf{x}^2 = \frac{\mathbf{y}^2}{|\mathbf{y}^2|} = \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0 \right)$$

is a unit eigenvector. It is easy to see that it is orthogonal to $\mathbf{x}^1$ (the scalar product of these vectors is zero). Finally, for $\lambda_3 = 1 - \sqrt{3}$ we find that

$$\mathbf{x}^3 = \left(\frac{1}{\sqrt{6+2\sqrt{3}}}, \frac{1}{\sqrt{6+2\sqrt{3}}}, -\frac{1+\sqrt{3}}{\sqrt{6+2\sqrt{3}}} \right)$$

is the third eigenvector.

The orthogonal matrix for passing over from the coordinates of the vector $\mathbf{x} = (x_1, x_2, x_3)$ in the system $(\mathbf{i}, \mathbf{j}, \mathbf{k})$ to the coordinates of the vector $\mathbf{x} = (\xi_1, \xi_2, \xi_3)$ in the system $(\mathbf{x}^1, \mathbf{x}^2, \mathbf{x}^3)$ has the form

$$\Lambda = \begin{vmatrix} x_1^{(1)} & x_2^{(1)} & x_3^{(1)} \\ x_1^{(2)} & x_2^{(2)} & x_3^{(2)} \\ x_1^{(3)} & x_2^{(3)} & x_3^{(3)} \end{vmatrix}$$

($|\Lambda| = 1$, $\Lambda^{-1}(\mathbf{i}) = \mathbf{x}^1$, $\Lambda^{-1}(\mathbf{j}) = \mathbf{x}^2$, $\Lambda^{-1}(\mathbf{k}) = \mathbf{x}^3$), i. e.

$$\left. \begin{aligned} x_1 &= x_1^{(1)}\xi_1 + x_1^{(2)}\xi_2 + x_1^{(3)}\xi_3, \\ x_2 &= x_2^{(1)}\xi_1 + x_2^{(2)}\xi_2 + x_2^{(3)}\xi_3, \\ x_3 &= x_3^{(1)}\xi_1 + x_3^{(2)}\xi_2 + x_3^{(3)}\xi_3. \end{aligned} \right\} \quad (6)$$

This transformation retains the orientation (since $|\Lambda| = 1 > 0$), i. e. the system $(\mathbf{x}^1, \mathbf{x}^2, \mathbf{x}^3)$ is oriented in the same way as the initial system $(\mathbf{i}, \mathbf{j}, \mathbf{k})$.

Substituting, instead of x_v , their values expressed by formulas (6) into our quadratic form, we get its canonical form

$$(1 + \sqrt{3})\xi_1^2 + (1 - \sqrt{3})\xi_2^2.$$

Now we would like to present a more detailed study of the equations and the corresponding surfaces of the above indicated eight types.

The *ellipsoid*

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1 \quad (a, b, c > 0). \quad (7)$$

For $a = b = c$ ellipsoid (7) turns into a sphere of radius a with centre at the origin, i.e. into the locus of points lying at a distance a from the origin.

The quantities a, b, c are called the *semiaxes of the ellipsoid*.

If in equation (7) we replace (simultaneously or separately) x by $-x$, y by $-y$, z by $-z$, then the equation remains unchanged. This shows that ellipsoid (7) is a surface symmetric both about the coordinate planes $x = 0$, $y = 0$, $z = 0$ and the origin. Therefore, it is sufficient to study the equation of ellipsoid (7) in the first octant (of the coordinate system), i.e. for $x \geq 0$, $y \geq 0$, $z \geq 0$. The portion of the ellipsoid located in the first octant is defined by an explicit equation, for instance,

$$z = c \sqrt{1 - \frac{x^2}{a^2} - \frac{y^2}{b^2}}, \quad x \geq 0, \quad y \geq 0, \quad \frac{x^2}{a^2} + \frac{y^2}{b^2} \leq 1.$$

For the sake of definiteness, we shall assume that $a \geq b \geq c$. The ellipsoid is a bounded surface. It is found inside a sphere of radius a with the origin as centre: the coordinates of any point (x, y, z) of the ellipsoid are related by the following inequality:

$$x^2 + y^2 + z^2 \leq a^2 \left(\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} \right) = a^2 \cdot 1 = a^2.$$

In order to form a more exact notion about the ellipsoid, let us intersect it by cutting planes parallel to the coordinate planes. For instance, cutting the ellipsoid by planes $z = h$ ($-c \leq h \leq c$), we shall intersect it in ellipses

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 - \frac{h^2}{c^2}$$

with the semiaxes

$$a \sqrt{1 - \frac{h^2}{c^2}}, \quad b \sqrt{1 - \frac{h^2}{c^2}}.$$

Hence it is seen that the greatest ellipse is yielded by the cutting plane $z = 0$. A similar result will be obtained by cutting the ellipsoid with the planes $x = h$ ($-a \leq h \leq a$), $y = h$ ($-b \leq h \leq b$).

Ellipsoid (7) has the form represented in Fig. 42.

The points $(\pm a, 0, 0)$, $(0, \pm b, 0)$, $(0, 0, \pm c)$ lie on ellipsoid (7) and are called its *vertices*.

If any two semiaxes are equal to each other, then ellipsoid (7) will be the *ellipsoid of revolution*, i.e. an ellipsoid generated by rotating an ellipse about one of the coordinate axes.

The *hyperboloid of one sheet*

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1 \quad (a, b, c > 0). \quad (8)$$

Judging by the form of equation (8), we conclude that the hyperboloid of one sheet is a surface symmetric both

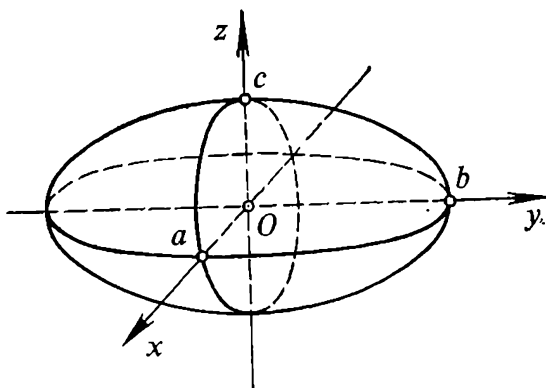


Fig. 42.

about the coordinate planes and the origin. The quantities a, b, c are called the *semiaxes of the hyperboloid of one sheet*. The points $(\pm a, 0, 0)$, $(0, \pm b, 0)$, lying on the surface of hyperboloid (8) are called the *vertices* of the hyperboloid of one sheet.

Cutting surface (8) by a plane $z = h$, we intersect it in an ellipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 + \frac{h^2}{c^2}$$

with the semiaxes

$$a \sqrt{1 + \frac{h^2}{c^2}}, \quad b \sqrt{1 + \frac{h^2}{c^2}}.$$

With h varying from $-\infty$ to $+\infty$, this ellipse describes surface (8).

Intersecting now surface (8) with a plane $x = h$ (or $y = h$), we shall cut off a hyperbola

$$\frac{y^2}{b^2} - \frac{z^2}{c^2} = 1 - \frac{h^2}{a^2} \left(\frac{x^2}{a^2} - \frac{z^2}{c^2} = 1 - \frac{h^2}{b^2} \right).$$

For $h = \pm a$ the first hyperbola decomposes into two straight lines $y = \pm \frac{b}{c} z$.

If $|h| \leq a$, then the real transverse axis of symmetry of the corresponding hyperbola is a straight line parallel to the y -axis, and for $|h| > a$ it is a straight line parallel to the z -axis.

The *real (transverse) axis of symmetry* is defined as one of the axes of symmetry which is intersected by the hyperbola.

If $a = b$, then planes $z = h$ will cut surface (8) in circles of radius $a \sqrt{1 + (h^2/c^2)}$. In this case surface (8) is generated by rotating the hyperbola $\frac{x^2}{a^2} - \frac{z^2}{c^2} = 1$ about the z -axis. The general view of a hyperboloid of one sheet is given in Fig. 43.

The *hyperboloid of two sheets*

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1 \quad (a, b, c > 0). \quad (9)$$

Since equation (9) contains only the squares of the variables, the given surface is symmetric both about the planes $x = 0$, $y = 0$, $z = 0$ and the origin.

Let us now write equation (9) in the general form

$$\frac{y^2}{b^2} + \frac{z^2}{c^2} = -1 + \frac{x^2}{a^2}. \quad (9')$$

Hence it is clear that, cutting surface (9') with a plane $x = h$ ($|h| \geq a$), we shall obtain an ellipse

$$\frac{y^2}{b^2} + \frac{z^2}{c^2} = -1 + \frac{h^2}{a^2}$$

with the semiaxes

$$b \sqrt{(h^2/a^2) - 1}, \quad c \sqrt{(h^2/a^2) - 1}.$$

For $|h| < a$ the number $(h^2/a^2) - 1 < 0$, and, therefore, there are no points of intersection of surface (9') and the plane $x = h$.

Cutting surface (9) with planes $z = h$ ($y = h$), we get hyperbolas

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 + \frac{h^2}{c^2} \left(\frac{x^2}{a^2} - \frac{z^2}{c^2} = 1 + \frac{h^2}{b^2} \right).$$

The points $(\pm a, 0, 0)$ lie on surface (9) and are called

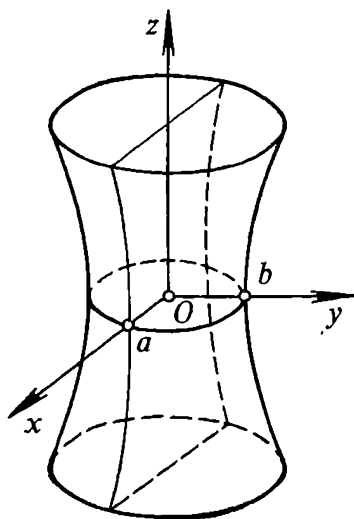


Fig. 43.

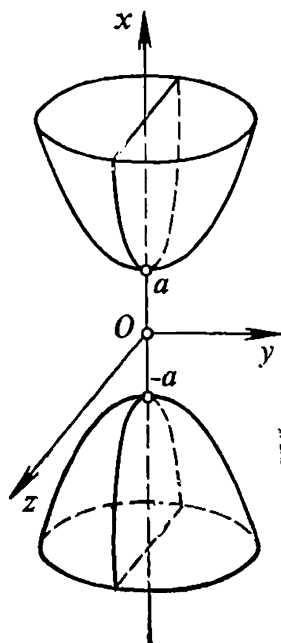


Fig. 44.

the *vertices of the parted (two-sheeted) hyperboloid*. Surface (9) is represented in Fig. 44.

The *elliptic paraboloid*

$$\frac{x^2}{p} + \frac{y^2}{q} = 2z \quad (p, q > 0). \quad (10)$$

Since equation (10) contains the squares of the variables x and y , we conclude that the given surface is sym-

metric about the coordinate planes $x = 0$, $y = 0$. Further, since we hold that $p, q > 0$, surface (10) is situated in the half space $z \geq 0$.

Cutting surface (10) with planes $z = h$ ($h \geq 0$), we shall intersect it in ellipses

$$\frac{x^2}{p} + \frac{y^2}{q} = 2$$

with the semiaxes

$$\sqrt{2ph}, \quad \sqrt{2qh}.$$

As h varies from zero to ∞ , these ellipses describe our surface (10).

Cutting surface (10) with planes $x = h$ (or $y = h$), we intersect it in parabolas

$$y^2 = 2q \left(z - \frac{h^2}{2p} \right) \quad \left(x^2 = 2p \left(z - \frac{h^2}{2q} \right) \right)$$

with a displaced vertex at the point $z = \frac{h^2}{2p}$ ($z = \frac{h^2}{2q}$).

For $p = q$ surface (10) will be a surface of revolution generated by rotating the parabola $x^2 = 2pz$ about the z -axis. In this case surface (10) is called the *paraboloid of revolution*.

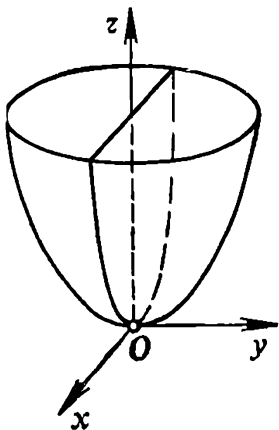


Fig. 45.

The point $(0, 0, 0)$ lies on surface (10) and is known as the *vertex of the elliptic paraboloid*. The elliptic paraboloid is represented in Fig. 45.

The *hyperbolic paraboloid*

$$\frac{x^2}{p} - \frac{y^2}{q} = 2z \quad (p, q > 0). \quad (11)$$

Judging by the form of equation (11), we conclude that this surface is symmetric about the planes $x = 0$, $y = 0$. Cutting surface (11)

with planes $z = h$, we obtain hyperbolas

$$\frac{x^2}{p} - \frac{y^2}{q} = 2h,$$

where for $h > 0$ the real axis of symmetry of the hyperbola will be parallel to the x -axis, and for $h < 0$ to the y -axis. For $h = 0$ we shall get a pair of intersecting straight lines.

When cutting surface (11) with planes $x = h$ or $y = h$, we shall get parabolas with their branches directed downwards or upwards:

$$-\frac{y^2}{q} = 2z - \frac{h^2}{p}, \quad \frac{x^2}{p} = 2z + \frac{h^2}{q}.$$

Surface (11) is depicted in Fig. 46.

The cone of the second order

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 0 \quad (a, b, c > 0). \quad (12)$$

It is clear that the given surface is symmetric both about the planes $x = 0$, $y = 0$, $z = 0$ and the origin.

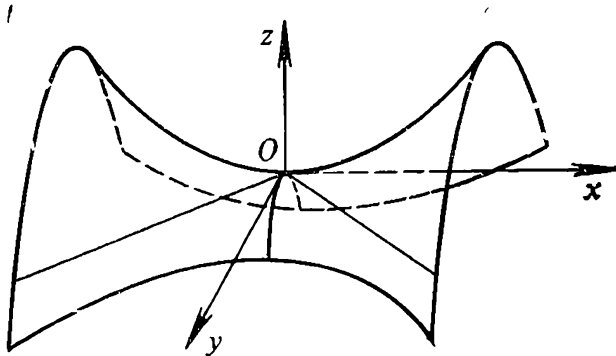


Fig. 46.

When intersecting surface (12) with planes $z = h$, we shall obtain ellipses

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = \frac{h^2}{c^2}$$

with the semiaxes $a | h | / c$ and $b | h | / c$.

And if surface (12) is cut by planes $x = h$ or $y = h$, then it will be intersected in hyperbolas

$$\frac{z^2}{c^2} - \frac{y^2}{b^2} = \frac{h^2}{a^2} \quad \left(\frac{z^2}{c^2} - \frac{x^2}{a^2} = \frac{h^2}{b^2} \right).$$

Returning to (11) and intersecting it with planes $y = hx$, we shall get a pair of intersecting lines

$$z = \pm cx \sqrt{(1/a^2) + (h^2/b^2)}.$$

The cone under consideration is shown in Fig. 47.

The point

$$x^2 + y^2 + z^2 = 0. \quad (13)$$

This equation is satisfied only by one point: $x = y = z = 0$.

Cylinders of the second order:

(a) The *elliptical cylinder*

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \quad (a, b > 0). \quad (14)$$

Equation (14) does not contain the variable z . On the plane xOy , equation (14) defines an ellipse with the semi-axes a and b . If a point (x, y) lies on this ellipse, then

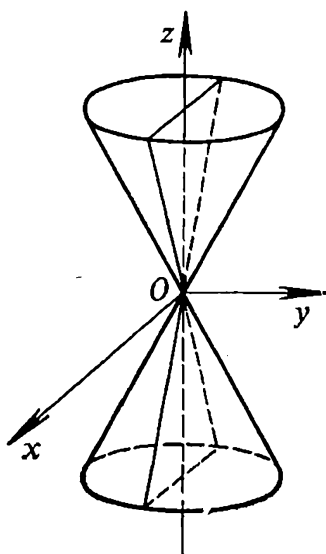


Fig. 47.

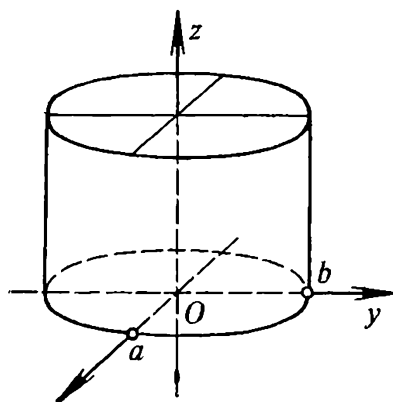


Fig. 48.

for any z the point (x, y, z) lies on surface (14). The set of such points is a surface described by a straight line parallel to the z -axis and intersecting the ellipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

in the plane xOy .

Ellipse (14) is called the *directrix* of the surface in question, and all possible positions of the generating line (generatrix)—the *elements*.

In general, a surface generated by a straight line which remains parallel to a certain given direction and intersects a given curve L is called *cylindrical*. Surface (14) is shown in Fig. 48.

(b) *Hyperbolic and parabolic cylinders*

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 \quad (a, b > 0), \quad (15)$$

$$y^2 = 2px \quad (p > 0). \quad (16)$$

In this case the directrices are a hyperbola and parabola, and the elements are straight lines parallel to the z -axis and passing through the hyperbola and parabola in the plane xOy .

Surfaces (15) and (16) are represented in Figs. 49 and 50, respectively.

(c) *Parallel and intersecting planes. A straight line*

$$a^2x^2 - b^2y^2 = 0 \quad (a, b > 0), \quad (17)$$

$$x^2 - a^2 = 0 \quad (a > 0), \quad (18)$$

$$z^2 = 0, \quad (19)$$

$$x^2 + y^2 = 0. \quad (20)$$

For surface (17) the directrices are straight lines

$$y = \pm \frac{a}{b} x.$$

Therefore, surface (17) is a pair of intersecting planes. The equations of both surfaces (18) and (19) are void of two coordinates each. Equation (18) in the plane xOy is a pair of straight lines $x = \pm a$.

If we shall take $x = \pm a$ and any y and z , then the points $(\pm a, y, z)$ will satisfy equation (18), therefore surface (18) is a pair of parallel planes.

Equation (19) describes the plane xOy , since this equation is satisfied by any points of the form $(x, y, 0)$ the entire set of which just constitute the plane xOy .

It is also possible to regard $z = 0$ as a directrix in one of the planes xOz or yOz , where the elements are straight

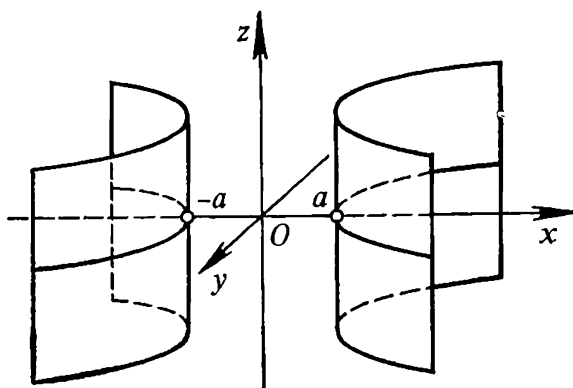


Fig. 49.

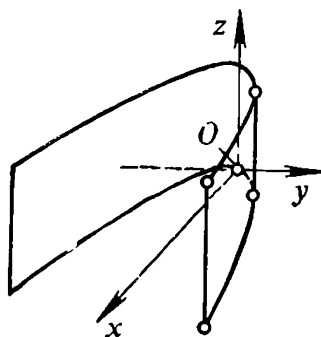


Fig. 50.

lines parallel either to the y - or x -axis and passing through the straight line $z = 0$.

Equation (20) is satisfied by any point with $x = y = 0$ and any z . Therefore, this equation defines a straight line, namely, the z -axis.

Ruled surfaces.

Some surfaces of the second order are generated by a moving straight line. Such are all cylindrical surfaces and a second-order cone. But there are other surfaces which are also generated by a moving straight line.

A surface that can be generated by a moving straight line is called a *ruled surface*, and the straight lines entirely lying in it are termed *rectilinear generators*.

The hyperboloid of one sheet and hyperbolic paraboloid belong to the ruled surfaces.

The equation of hyperboloid of one sheet (8) can be written in the form

$$\frac{x^2}{a^2} - \frac{z^2}{c^2} = 1 - \frac{y^2}{b^2}$$

or, factorizing both sides of the equation, we get

$$\left(\frac{x}{a} + \frac{z}{c}\right) \left(\frac{x}{a} - \frac{z}{c}\right) = \left(1 + \frac{y}{b}\right) \left(1 - \frac{y}{b}\right). \quad (21)$$

Let us compose a system of the first degree:

$$\left. \begin{aligned} \frac{x}{a} + \frac{z}{c} &= k \left(1 + \frac{y}{b}\right), \\ \frac{x}{a} - \frac{z}{c} &= \frac{1}{k} \left(1 - \frac{y}{b}\right), \end{aligned} \right\} \quad (22)$$

where k is an arbitrary parameter.

For a definite value of the parameter k we obtain a straight line, thus, as k varies, we get a *family of straight lines*. Multiplying equations (22) termwise, we shall obtain equation (21) of our surface. Therefore, any point (x, y, z) , satisfying system (22), is located on surface (21). Consequently, each line belonging to family (22) lies entirely on the surface of the hyperboloid of one sheet.

Quite analogously, the system

$$\left. \begin{aligned} \frac{x}{a} + \frac{z}{c} &= l \left(1 - \frac{y}{b}\right), \\ \frac{x}{a} - \frac{z}{c} &= \frac{1}{l} \left(1 + \frac{y}{b}\right), \end{aligned} \right\} \quad (23)$$

where l is a parameter, also defines a family of straight lines, which is different from family (22) and belongs to surface (21).

Through every point of hyperboloid (21) there passes one straight line of either family, generally speaking, for different values of the parameters k and l (Fig. 51). For instance, through the point $\left(\sqrt{\frac{3}{2}}a, \frac{\sqrt{2}}{2}b, c\right)$ of surface (21)

there passes a straight line (22) for $k = (2 + \sqrt{6})/(2 + \sqrt{2})$ and a straight line (23) for $l = (2 + \sqrt{6})/(2 - \sqrt{2})$.

The ruled-surface nature of a hyperboloid of one sheet was utilized by the Soviet engineer V. G. Shukhov (1853-

1939) in the construction of what is called the "Shukhov Tower" of Moscow (in the Shabolovka street) which for years was used as the Moscow radio and television tower. It was constructed out of steel strips arranged along rectilinear generatrices of a hyperboloid of one sheet. The strips were riveted together at the points of intersection of the two systems of generators. Shukhov's structure possesses high strength, though a relatively small amount of material was used in the construction. Various tall towers have been constructed using Shukhov's method.

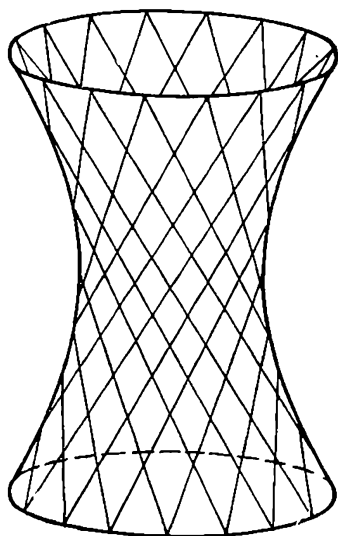


Fig. 51.

It can be easily verified that the two families of straight lines:

$$\left. \begin{aligned} \frac{x}{\sqrt{p}} + \frac{y}{\sqrt{q}} &= 2kz, \\ \frac{x}{\sqrt{p}} - \frac{y}{\sqrt{q}} &= \frac{1}{k}, \end{aligned} \right\} \quad (24)$$

$$\left. \begin{aligned} \frac{x}{\sqrt{p}} + \frac{y}{\sqrt{q}} &= \frac{1}{l}, \\ \frac{x}{\sqrt{p}} - \frac{y}{\sqrt{q}} &= 2lz \end{aligned} \right\} \quad (25)$$

generate the surface of hyperbolic paraboloid (11).

The straight lines belonging to families (24) and (25) lie on this surface and, conversely, any point of this surface is the intersection of a certain straight line of family (24) and a line belonging to family (25).

Sec. 26. The General Theory of a Second-order Surface in Three-dimensional Space

Let there be given a surface of the second order

$$\sum_{k=1}^3 \sum_{l=1}^3 a_{kl} x_k x_l + 2 \sum_{l=1}^3 A_l x_l + B = 0, \quad (1)$$

where $a_{kl} = a_{lk}$, A_l , B are constant numbers—the coefficients of the equation.

The preceding section described eight forms ((1) to (8)) (particular cases) of equation (1) and it was noted there that for every given equation (1), provided that it defines not an imaginary surface, it is possible to find a rectilinear coordinate system in which this equation has one of the indicated forms.

This statement is proved below.

To begin with, we consider the quadratic form standing in the left-hand side of equation (1).

On the basis of Theorem 2 of Sec. 22, this form can be reduced, with the aid of the orthogonal transformation

$$x_i = \sum_{s=1}^3 \beta_{is} x'_s \quad (i = 1, 2, 3), \quad (2)$$

to the following form:

$$\sum_{k=1}^3 \sum_{l=1}^3 a_{kl} x_k x_l = \lambda_1 x_1'^2 + \lambda_2 x_2'^2 + \lambda_3 x_3'^2,$$

where $\lambda_1, \lambda_2, \lambda_3$ are definite real numbers.

We should like to underline that equalities (2) define the transformation of the initial rectangular coordinate system (x_1, x_2, x_3) to some other rectangular system (x'_1, x'_2, x'_3) . The point, whose coordinates in the initial system is (x_1, x_2, x_3) , in the new system has the coordinates (x'_1, x'_2, x'_3) obtained by inverting operation (2).

In the new rectangular system our surface has, obviously, the following equation:

$$\lambda_1 x_1'^2 + \lambda_2 x_2'^2 + \lambda_3 x_3'^2 + 2 \sum_{k=1}^3 A'_k x'_k + B = 0, \quad (3)$$

where A'_k are some constant numbers.

Let us first consider the case when all the three numbers are different from zero ($\lambda_i \neq 0, i = 1, 2, 3$).

In this case we transfer the coordinate system x'_1, x'_2, x'_3 so that its origin goes into the point (a_1, a_2, a_3) ; we then obtain a second rectangular coordinate system (ξ_1, ξ_2, ξ_3) , where

$$x'_i = a_i + \xi_i \quad (i = 1, 2, 3).$$

In it the equation of our surface has the form

$$\lambda_1 (a_1 + \xi_1)^2 + \lambda_2 (a_2 + \xi_2)^2 + \lambda_3 (a_3 + \xi_3)^2 + 2 \sum_{k=1}^3 A'_k (a_k + \xi_k) + B = 0$$

or

$$\lambda_1 \xi_1^2 + \lambda_2 \xi_2^2 + \lambda_3 \xi_3^2 + 2 \sum_{l=1}^3 (a_l \lambda_l + A'_l) \xi_l + B_1 = 0,$$

where B_1 is some constant. Setting

$$a_l = -\frac{A'_l}{\lambda_l} \quad (l=1, 2, 3),$$

we simplify this equation:

$$\lambda_1 \xi_1^2 + \lambda_2 \xi_2^2 + \lambda_3 \xi_3^2 = -B_1. \quad (4)$$

Suppose the numbers $\lambda_1, \lambda_2, \lambda_3$ are of the same sign.

If in this case $B_1 = 0$, then equation (4) is satisfied only by *one point*, namely, by the zero point $(0, 0, 0)$ (see (7) in the preceding section).

If $B_1 \neq 0$ and has the same sign as the numbers $\lambda_1, \lambda_2, \lambda_3$, then, obviously, there are no points with real coordinates which satisfy equation (4). In this case *surface (1) is imaginary*.

And if $B_1 \neq 0$ and has the sign opposite to that of the numbers $\lambda_1, \lambda_2, \lambda_3$, then equation (4) can be written in the form

$$-\frac{\xi_1^2}{\frac{B_1}{\lambda_1}} + -\frac{\xi_2^2}{\frac{B_1}{\lambda_2}} + -\frac{\xi_3^2}{\frac{B_1}{\lambda_3}} = 1 \quad (5)$$

or, setting

$$a^2 = -\frac{B_1}{\lambda_1}, \quad b^2 = -\frac{B_1}{\lambda_2}, \quad c^2 = -\frac{B_1}{\lambda_3},$$

in the form

$$\frac{\xi_1^2}{a^2} + \frac{\xi_2^2}{b^2} + \frac{\xi_3^2}{c^2} = 1 \quad (a, b, c > 0).$$

Hence, surface (1) is an *ellipsoid* (see (1) in the preceding section).

Let now two of the numbers $\lambda_1, \lambda_2, \lambda_3$ be of the same sign and the third have an opposite sign.

If $B_1 = 0$, then in equation (4) we may regard that $\lambda_1 > 0, \lambda_2 > 0, \lambda_3 < 0$, multiplying (4), if necessary, by -1 and interchanging, if necessary, ξ_j . Then, making

$$\lambda_1 = \frac{1}{a^2}, \quad \lambda_2 = \frac{1}{b^2}, \quad \lambda_3 = -\frac{1}{c^2} \quad (a, b, c > 0),$$

we shall get the equation of a *cone* (see (6) in the preceding section)

$$\frac{\xi_1^2}{a^2} + \frac{\xi_2^2}{b^2} - \frac{\xi_3^2}{c^2} = 0.$$

For $B_1 \neq 0$ we take advantage of formula (5) once again. We distinguish here two essentially different cases:

$$1. \frac{\xi_1^2}{a^2} + \frac{\xi_2^2}{b^2} - \frac{\xi_3^2}{c^2} = 1 \quad (\text{hyperboloid of one sheet, see (2)}$$

in Sec. 25),

$$2. \frac{\xi_1^2}{a^2} - \frac{\xi_2^2}{b^2} - \frac{\xi_3^2}{c^2} = 1 \quad (\text{hyperboloid of two sheets, see (3)}$$

in Sec. 25).

The remaining cases are reduced to these two by the corresponding replacement of the coordinates ξ_i .

Let now $\Delta = \lambda_1 \lambda_2 \lambda_3 = 0$. Then at least one of the numbers $\lambda_1, \lambda_2, \lambda_3$ is equal to zero. We shall regard that $\lambda_3 = 0$, interchanging, if necessary, ξ_i .

So, let $\lambda_3 = 0$. We first single out the case when $A'_3 = 0$. Then equation (3) has the form

$$\lambda_1 x_1'^2 + \lambda_2 x_2'^2 + 2(A'_1 x_1' + A'_2 x_2') + B = 0,$$

where $\lambda_1, \lambda_2, A'_1, A'_2, B$ are any numbers.

In the plane (x'_1, x'_2) this is the general equation of a second-order curve. In space (x'_1, x'_2, x'_3) , it is the *equation of a cylindrical surface* (see (8) in the preceding section) passing through a *plane second-order curve* with the generatrix parallel to the axis x'_3 (see (8) in the preceding section).

Further, we shall always hold that $A'_3 \neq 0$ and $\lambda_3 = 0$, and then equation (3) has the form

$$(\lambda_1 x_1'^2 + 2A'_1 x_1') + (\lambda_2 x_2'^2 + 2A'_2 x_2') + 2A'_3 x_3' + B = 0. \quad (3')$$

Here, we distinguish the following essentially different cases: (a) $\lambda_1, \lambda_2 > 0$; (b) $\lambda_1 > 0, \lambda_2 < 0$; (c) $\lambda_1 > 0, \lambda_2 = 0$. Other cases are reduced to these by replacing the coordinates or multiplying by -1 .

Consider Case (a) $\lambda_1, \lambda_2 > 0$. Take the factors λ_1 and λ_2 out of the parentheses and complete the expressions between these parentheses to perfect squares. Then, taking into account that $\lambda_1, \lambda_2, A'_3$ are different from zero, we shall get

$$\lambda_1 (x_1' + \alpha)^2 + \lambda_2 (x_2' + \beta)^2 + 2A'_3 (x_3' + \gamma) = 0,$$

where α, β, γ are relevant numbers. Setting

$$\xi = x_1' + \alpha, \quad \eta = x_2' + \beta, \quad \zeta = x_3' + \gamma,$$

we get

$$\lambda_1 \xi^2 + \lambda_2 \eta^2 = -2A'_3 \zeta$$

or

$$\frac{\xi^2}{-\frac{A'_3}{\lambda_1}} + \frac{\eta^2}{-\frac{A'_3}{\lambda_2}} = 2\zeta. \quad (6)$$

If $-A'_3 > 0$, then equation (6) has the form

$$\frac{\xi^2}{p} + \frac{\eta^2}{q} = 2\zeta \quad (\text{elliptic paraboloid, see (4) in Sec. 25}) \quad (7)$$

$$(p, q > 0).$$

And if $-A'_3 < 0$, then replacing ζ by $-\zeta$, we get once again the equation of form (7), i.e. an *elliptic paraboloid*.

Consider now Case (b) $\lambda_1 > 0, \lambda_2 < 0, (A'_3 \neq 0)$. We take advantage of equation (6). If $A'_3 < 0$, then this equation is written in the form

$$\frac{\xi^2}{p} - \frac{\eta^2}{q} = 2\zeta \quad (\text{hyperbolic paraboloid, see (5) in Sec. 25}) \quad (8)$$

$$(p, q > 0).$$

And if $A'_3 > 0$, then, on interchanging ξ and η we shall get again the equation of form (8), i.e. a *hyperbolic paraboloid*.

Finally, we pass over to Case (c) $\lambda_1 > 0$, $\lambda_2 = 0$ ($A'_3 \neq 0$). Then equation (3') is reduced to the following:

$$(\lambda_1 x_1'^2 + 2A'_1 x_1) + 2(A'_2 x_2' + A'_3 x_3') + B = 0.$$

Factoring λ_1 out of the first parentheses and supplementing the expression between these parentheses to a perfect square (taking into consideration that $\lambda_1, A'_3 \neq 0$), we obtain

$$\lambda_1 (x_1' + \alpha)^2 + 2A'_2 x_2' + 2A'_3 (x_3' + \beta) = 0,$$

where α, β are some numbers. On substituting

$$\xi = x_1' + \alpha, \quad \eta = x_2', \quad \zeta = x_3' + \beta,$$

this equation will turn into the following:

$$\lambda_1 \xi^2 + 2(A'_2 \eta + A'_3 \zeta) = 0. \quad (9)$$

Consider in the plane (η, ζ) the vector $\omega = (A'_2, A'_3)$. Let us write it in the form

$$(A'_2, A'_3) = \rho (\cos \alpha, \sin \alpha),$$

where $\rho > 0$ is the length of ω , and $(\cos \alpha, \sin \alpha)$ is a unit vector in the direction of ω . The second vector $(\sin \alpha, -\cos \alpha)$ is also a unit vector, it is perpendicular to the first one.

Let us introduce in the plane (η, ζ) the following orthogonal transformation:

$$u = \eta \cos \alpha + \zeta \sin \alpha, \quad v = \eta \sin \alpha - \zeta \cos \alpha.$$

By means of this transformation we replace the rectangular coordinate system (η, ζ) by the rectangular coordinate system (u, v) , and the rectangular coordinate system (ξ, η, ζ) by the rectangular coordinate system (ξ, u, v) . As a result, equation (9), which can be written as

$$\lambda_1 \xi^2 + 2\rho (\eta \cos \alpha + \zeta \sin \alpha) = 0,$$

takes the following form:

$$\lambda_1 \xi^2 + 2pu = 0,$$

or, substituting $-u$ for u ,

$$\xi^2 = 2pu \quad (p = \rho/\lambda_1 > 0),$$

or, finally, interchanging ξ and u ,

$$u^2 = 2p\xi \quad (p > 0),$$

i.e. we have obtained the *equation of a parabolic cylinder* (in the rectangular coordinates (ξ, u, v)).

Thus, we have considered all possible cases for equation (1) and in each of them have found the rectangular coordinate system in which equation (1) has one of the forms listed in Sec. 25 ((1) through (8)). The statement has been proved.

SUBJECT INDEX

- Abscissa, 44
- Angle(s)
 - between n -dimensional vectors, 52
 - between two planes, 71
- Buniakovski's inequality, 51
- Cone of the second order, 175
- Coordinate axes, 44
- Coordinate system
 - left-handed, 80
 - right-handed, 80
 - in space, rectangular, 44
- Coordinates of a point, 48
- Cylinders of the second order, 176
 - elliptic, 176
 - hyperbolic, 177
 - parabolic, 177
- Determinant(s)
 - columns of, 7
 - diagonals of, 8
 - leading, 8
 - principal, 8
 - secondary, 8
 - of the n th order, 11
 - properties of, 7
 - rows of, 7
 - second-order, 7
 - term of, 9, 11
 - third-order, 8
 - Vandermonde (or power), 16
- Distance from point to plane, 72
- Element(s)
 - cofactor of, 13
 - of determinant, 8
 - minor of, 13
- Ellipse, 147
 - equation of, 147
 - foci of, 148
- Ellipsoid, 170
 - of revolution, 171
 - semiaxes of, 170
 - vertices of, 171
- Elliptic cylinder, 176
- Equation(s)
 - of a plane
 - general, 67
 - intersept form of, 69
 - normal form of, 67
 - in a vector form, 67
 - of a straight line
 - in canonical form, 75
 - general, 58
 - intercept form, 60
 - normal form of, 63
 - passing through a point, 60
- Foci
 - of ellipse, 148
 - of hyperbola, 151
- Focus (*see* Foci)

- Gauss method, 37
- Hyperbola(s), 151
 asymptotes of, 152
 foci of, 151
 vertices of, 152
- Hyperboloid
 of one sheet, 171
 of two sheets, 172
- Imaginary surface, 162
- Line segment, 54
- Matrix(ces), 19
 augmented, 24
 columns of, 19
 conjugate, 20
 elements of, 19
 product of, 98
 rank of, 20
 rows of, 19
 square, 19
 sum of, 20
 transpose of, 20
 unit, 99
- Method of elimination of unknowns, 33
- Minkowski's inequality, 52
- Normalizing factor, 63
- Operator(s), 96
 characteristic equation of, 143
 eigenvalue of, 135
 eigenvector of, 135
 inverse of, 99
 linear, 97, 132
 self-adjoint, 133
 unit, 99
- Ordinate, 44
- Origin (of coordinates), 44
- Parabola(s), 156
 directrix of, 156
 focus of, 156
- Parabolic cylinder, 177
- Paraboloid
 elliptic, 173
 hyperbolic, 174
 of revolution, 174
- Permutation (s), 10
 even, 10
 of digits, 10
 main, 10
 odd, 10
- Perpendicularity
 of planes, condition for, 72
 of straight lines, condition for, 62
- Quadratic form, 133
 canonical form of, 145
 strictly negative, 141
 strictly positive, 140
- Radius vector, 44
- Rectilinear generator, 178
- Ruled surface, 178
- Scalar product of two vectors, 43
- Space,
 Euclidean, 50
 n -dimensional, 48
 three-dimensional, 48
- Straight line(s)
 general equation of, 58
 intercept form of equation of, 60
 normal form of equation of, 63

-
- passing through a point, equation of, 60
 - slope of, 58
 - Subspace, 119
 - linear, 119
 - m -dimensional, 119
 - one-dimensional, 121
 - orthogonal, 121
 - Sylvester's theorem, 141
 - System(s)
 - coefficients of, 22
 - of equations, 22
 - homogeneous, 24
 - solution of, 22, 24
 - of vectors
 - linearly dependent, 89
 - linearly independent, 89
 - Three-dimensional space, 47
 - Transposition(s)
 - number of, 10
 - of two elements, 10
 - Triple scalar product of three vectors, 87
 - Unit vector, 108
 - normal, 108
 - Vector(s), 21, 39
 - basis, 146
 - collinear, 39
 - components of, 21, 48
 - coplanar, 80
 - length of, 39
 - n -dimensional, 21
 - null, 40
 - projection of, 40
 - sum of, 49
 - Vector product of two vectors, 81
 - Z-coordinate, 44

Higher Mathematics

Ya. S. BUGROV, S. M. NIKOLSKY

Other Books in the Series

Differential and Integral Calculus

Differential Equations

Multiple Integrals

Series

Theory of Functions
of a Complex Variable

A Collection of Problems

MIR PUBLISHERS MOSCOW